

RICE UNIVERSITY

**An Empirical Analysis of Internet Use on Smartphones:  
Characterizing Visit Patterns and User Differences**

by

**Chad C. Tossell**

A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

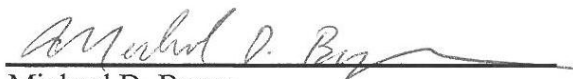
**Doctor of Philosophy**

APPROVED, THESIS COMMITTEE



---

Philip Kortum, Chair  
Professor-in-the-Practice, Psychology



---

Michael D. Byrne  
Associate Professor, Psychology



---

David M. Lane  
Associate Professor, Psychology



---

Lin Zhong  
Associate Professor, Electrical and  
Computer Engineering

HOUSTON, TEXAS  
May 2012

## ABSTRACT

### **An Empirical Analysis of Internet Use on Smartphones: Characterizing Visit Patterns and User Differences**

by

**Chad C. Tossell**

The original vision of ubiquitous computing was for computers to assist humans by providing subtle and fitting technologies in every environment. The iPhone and similar smartphones have provided continuous access to the internet to this end. In the current thesis, my goal was to characterize how the internet is used on smartphones to better understand what users do with technology away from the desktop. Naturalistic and longitudinal data were collected from iPhone users in the wild and analyzed to develop this understanding. Since there are two general ways to access the internet on smartphones—via native applications and a web browser—I describe usage patterns through each along with the influence of experience, the nature of the task and physical locations where smartphones were used on these patterns.

The results reveal differences between technologies (the PC and the smartphone), platforms (native applications and the mobile browser), and users in how the internet was accessed. Findings indicate that longitudinal use of web browsers decreased sharply with time in favor of native application use, web page revisitation through browsers occurred very infrequently (approximately 25% of URLs are revisited by each user), bookmarks were used sparingly to access web content, physical location visitation followed patterns similar to virtual visitation on the internet, and Zipf distributions characterize mobile

internet use. The web browser was not as central to smartphone use compared to the PC, but afforded certain types of activities such as searching and ad hoc browsing. In addition, users systematically differed from each other in how they accessed the internet suggesting different ways to support a wider spectrum of smartphone users.

## Acknowledgments

This dissertation would not have been possible without the support of my board, colleagues, friends and family. I would like to recognize these individuals here for making my graduate school experience so rewarding.

First, I am indebted to my advisor, Dr. Philip Kortum, for his mentoring, patience and instruction. I benefited greatly from the research we conducted together and, under his guidance, have learned a great many lessons that apply both to the field of human factors and beyond. Dr. Kortum also recognized my unique situation being in the military and ensured my requirements were being met at every stage. It has been a wonderful experience working for him and I look forward to our future collaborations.

Second, it has been an honor to work with the other members of my thesis committee, Dr. Michael Byrne, Dr. David Lane and Dr. Lin Zhong, over the past few years. I have grown tremendously as a scientist because of the training I received from each of them both in and out of the classroom.

Third, I have been privileged to collaborate with a number of colleagues in this research. In particular, Clayton Shepard has been the lead engineer in the design and implementation of the logger used to collect data for this research. Similarly, Ahmad Rahmati has provided great technical insight and helped shape the overall study. Both have been instrumental in the completion of this thesis. I also thank Jeff Zemla for his assistance on various aspects of this project including an excellent review of a previous version of this paper. Several undergraduate researchers helped process and reduce the

enormous data set studies in this dissertation. I thank Amy Buxbaum, Beth Herlin, Louis Fornage, Michael Padon and Wen Xing for their tremendous efforts in these tasks.

I would be remiss if I did not acknowledge several others that have supported me in recent years. I thank Dr. Roger Schvaneveldt, Jason Parker, Dr. Terence Andre, David Shiflet and Greg Parker for their friendship, leadership and influence in all aspects of life. Additionally, I appreciate the consistent support of my parents throughout the years.

Finally, I am grateful to my wife and our six children. The former has been a truly exceptional wife in ways too numerous to mention in this section. Her encouragement inspired me to successfully complete this work and it has been an honor to be her husband. I would not be who I am without her influence. My children have also given me such joy and motivation to succeed in this research. I thank my wife and my children for their patience in letting me complete this work and I dedicate this dissertation to each of them.

# Contents

<b>Acknowledgments.....</b>	<b>iv</b>
<b>Contents .....</b>	<b>vi</b>
<b>List of Figures .....</b>	<b>viii</b>
<b>List of Tables .....</b>	<b>x</b>
<b>List of Equations .....</b>	<b>xi</b>
<b>INTRODUCTION .....</b>	<b>1</b>
1.1. Problem Statement .....	2
1.2. Goals .....	4
1.3. Approach .....	7
1.4. Overview.....	7
<b>USING THE WEB ON PERSONAL COMPUTERS .....</b>	<b>9</b>
<b>SMARTPHONES: CHANGING HOW THE INTERNET IS USED.....</b>	<b>15</b>
3.1. Two Platforms to Access the Internet on Smartphones .....	15
3.2. Mobile Usability.....	17
3.3. Context and the Nature of Smartphone Interactions .....	20
3.4. Chapter Summary and Look Ahead.....	25
<b>METHODOLOGY .....</b>	<b>27</b>
4.1. Overview.....	27
4.2. Data collection.....	31
4.2.1. Participants .....	31
4.2.2. Materials .....	31
4.2.3. Procedure .....	34
4.3. Data Analysis .....	36
4.3.1. Preprocessing the Data.....	36
4.3.2. Primary Measures of Interest.....	38
4.3.2.1. Recurrent Behaviors.....	39
4.3.2.2. EntroDiversity – Intra-user Variability in iPhone Use .....	40
4.3.2.3. NIA-to-Site Index .....	42
4.4. Strengths and Limitations of the Methodology .....	43

<b>RESULTS .....</b>	<b>47</b>
5.1. Web Browser Visitation Patterns .....	47
5.1.1. Revisiting Pages and Sites on the Mobile Web .....	48
5.1.1.1. The Influence of the Task on Revisiting Browser Content .....	53
5.1.1.2. Searching on the Mobile Web .....	56
5.2. Native Internet Application Visitation Patterns .....	58
5.3. NIA and Site Visits: Temporal and Contextual Patterns .....	62
5.3.1. Behavior Rates and Indices .....	65
5.3.2. Physical and Virtual (Internet) Revisitation Patterns .....	70
5.4. Models of Physical and Virtual Visitation .....	72
5.5. Exploring User Diversity .....	77
<b>DISCUSSION .....</b>	<b>81</b>
6.1. Characterizing Web Use on Smartphones .....	81
6.2. NIAs and Browsers: NIAs Dominate, But the Web is Not Dead .....	83
6.3. User Differences in Mobile Internet Use .....	89
6.4. Design Implications .....	92
6.5. Limitations .....	99
6.6. Conclusion .....	100
<b>REFERENCES .....</b>	<b>103</b>

## List of Figures

<b>Figure 1. Hypothetical day of computing without a smartphone (top) and with a smartphone (bottom) based on survey data (Ericcson, 2011). Thinner rectangles represent smartphone internet visits whereas thicker rectangles represent PC usage.</b>	<b>22</b>
<b>Figure 2. iPhone 3GS displaying the first springboard page with native applications that come standard on the device.</b>	<b>32</b>
<b>Figure 3. Implementation of logging methodology (from Shepard, Rahmati, Tossell, Zhong &amp; Kortum, 2010).</b>	<b>35</b>
<b>Figure 4. Three studies (1-current study, 2-Jonssen et al., 2002, 3-Oulasvirta et al., 2007) that employed smartphone logging and their approximate placements on each of the nine considerations.</b>	<b>46</b>
<b>Figure 5. Diurnal patterns of web browser usage.</b>	<b>48</b>
<b>Figure 6. URL vocabulary by total URL visits over the first 600 visits.</b>	<b>52</b>
<b>Figure 7. Distribution of page revisits within domains for each type of site (after Obendorf et al., 2007).</b>	<b>54</b>
<b>Figure 8. Screen capture of the bookmarks window on the Safari browser.</b>	<b>55</b>
<b>Figure 9. Percentage of search sessions by the number of queries issued.</b>	<b>57</b>
<b>Figure 10. Number of native applications installed by genre. The inset plots the percentage of NIA uninstalled by the same genre distinctions.</b>	<b>60</b>
<b>Figure 11. Cumulative percentage of native internet application visits.</b>	<b>61</b>
<b>Figure 12. Number of (a) total visits, (b) unique content and new content accessed through NIAs and the browser by month. Geometric means are used in these figures to reduce the influence of outliers within each month on the summary statistics.</b>	<b>63</b>
<b>Figure 13. Box plot of EntroDiversity scores by month.</b>	<b>66</b>
<b>Figure 14. Box plot of the proportion of NIAs used to total NIAs within each user's vocabulary by month.</b>	<b>67</b>
<b>Figure 15. The percentage of browsing sessions that included Google over time.</b>	<b>68</b>



<b>Figure 16. Scatter plot of total internet visits by NIA-to-site indices .....</b>	<b>69</b>
<b>Figure 17. Percentages of resources and locations by their number of visits.....</b>	<b>71</b>
<b>Figure 18. Top 10 sites, NIAs and locations for each user visited within each month. .....</b>	<b>72</b>
<b>Figure 19. Fitted Zipf distributions for ranked NIA, site, and location visits (left column) alongside the corresponding mean square error (MSE) of users' actual distributions against the Zipf distributions. The insets within the latter are log-log plots. ....</b>	<b>75</b>
<b>Figure 20. An example of an enhanced browser bar that facilitates easier access to top sites. Menus can be displayed for sites with high page revisits (e.g., institutional sites) to the sub-level pages that are frequently accessed. ....</b>	<b>94</b>
<b>Figure 21. An example of a personalizable "Web Springboard." The Opera mini browser provides users easy access to top sites from its home page. A similar approach could be used for a springboard designated for browser use. ....</b>	<b>95</b>
<b>Figure 22. Examples of intelligent springboards that leverage users' history of application installs (left) and location information (right).....</b>	<b>96</b>
<b>Figure 23. An example of more personalizable springboards for enhanced identification of NIAs.....</b>	<b>99</b>

## List of Tables

<b>Table 1. Nine considerations used in the design of this study.....</b>	<b>30</b>
<b>Table 2. Pre-installed applications for the iPhone 3GS.....</b>	<b>32</b>
<b>Table 3. Usage data collected by the logger.....</b>	<b>33</b>
<b>Table 4. Sample of NIAs analyzed along with several not studied.....</b>	<b>37</b>
<b>Table 5. Two web browser sessions recorded from Subject 10. ....</b>	<b>38</b>
<b>Table 6. Three hypothetical examples of users and their associated EntroDiversity indices. Only two applications are used for clarity alongside (visit counts).....</b>	<b>42</b>
<b>Table 7. Comparison of primary revisitation studies with the current study.....</b>	<b>49</b>
<b>Table 8. Frequency and percentage of URL Visits by Type. ....</b>	<b>53</b>
<b>Table 9. Comparisons between PC and smartphone searching behaviors.....</b>	<b>57</b>
<b>Table 10. Summary and variance statistics for native internet application use. ....</b>	<b>59</b>

## List of Equations

<b>Equation 1. Page Revisitation Rate .....</b>	<b>39</b>
<b>Equation 2. Site Revisitation Rate .....</b>	<b>39</b>
<b>Equation 3. NIA Revisitation Rate .....</b>	<b>40</b>
<b>Equation 4. Location Revisitation Rate .....</b>	<b>40</b>
<b>Equation 5. EntroDiversity Index .....</b>	<b>41</b>
<b>Equation 6. NIA-to-Site Index .....</b>	<b>43</b>
<b>Equation 7. Power Function.....</b>	<b>73</b>
<b>Equation 8. Logarithmic Transform of Power Function .....</b>	<b>73</b>
<b>Equation 9. Zipf Function .....</b>	<b>74</b>
<b>Equation 10. Logarithmic Transform of Zipf Function.....</b>	<b>74</b>

## Chapter 1

### INTRODUCTION

The World Wide Web (web) is a well-known resource for information, communication, and entertainment used by over two billion people worldwide (Internet World Statistics, 2011). This ubiquity has led to a large amount of human-computer interaction (HCI) research examining how it is used. Most work has focused on behaviors associated with the personal computer (PC). For instance, studies of web use on this technology have developed an understanding of user goals (Kellar, Watters & Shepherd, 2007), browsing strategies (Catledge & Pitkow, 1995), tasks (Byrne, John, Wehrle & Crow, 1999), search behaviors (Kamvar, Kellar, Pater & Xu, 2009; White & Drucker, 2007), information foraging (Pirolli & Card, 1995), revisitation of websites (Tauscher & Greenberg, 1997), and differences between groups (e.g., novice-expert; Holscher & Strube, 2000) among other efforts. A number of PC-based web studies logged users interacting with the web in more naturalistic settings to establish empirical patterns of behaviors. In developing models of internet use, these characterizations are rare, but the most informative (Pitkow, 1998). Results from several logs-based analyses informed the development of theories and models of web use on stationary computers, such as link selection based on information scent (Pirolli & Fu, 2003) and spread-activation models of predictive retrieval of web information (Pitkow, 1998).

Now, smartphones are becoming pervasive and making the web accessible in most personal and professional environments (Matthews, Pierce & Tang, 2009). These technologies are being adopted at higher rates than any other technology in history (Eagle, 2005). The International Telecommunications Union (ITU) reported that over 1.2

billion phones being used around the world can currently access the internet (2012). Estimates suggest that by the end of 2012, the number of activated iPhone and Android devices alone will exceed the one billion mark (Newark-French, 2012). Additional projections suggest that by the year 2015, there will be more mobile internet users than PC internet users (Hepburn, 2011). There is even some evidence that smartphone users spend more time using the internet on smartphones compared to their PCs (Newark-French, 2011). Clearly, we are at a historical shift in HCI where smartphones are changing the face of how the internet is accessed around the world.

### **1.1. Problem Statement**

Although there is a substantial amount of research focused on how the internet is used on PCs, there are several reasons why these studies may not apply to web interactions on smartphones. First, internet resources can be accessed on smartphones through two platforms. The first platform, the web browser, allows users to visit unique resource locators (URLs) similar to how they are visited on PCs. Many of the sites visited on the mobile web automatically display a version of the URL that is optimized for small screens. The second way to access the internet, more central to smartphone use (Kim, 2012), is through native applications (apps). These capabilities provide users access to mobile content designed specifically for their type of smartphones. Since users spend more time on the latter platform (Kim, 2012), it is unclear how both browsers and native applications on the same device are used to access the internet.

Second, users interact differently with smartphones compared to PCs. Navigating the web on a PC generally requires a keyboard and mouse. Current smartphones, led by

the iPhone, now allow users to traverse the internet on smaller touch screens. Gesture-based interactions afford actions such as scrolling, zooming, and moving to a subsequent page instead of points and clicks with a mouse. These capabilities have improved the usability of web interactions with smartphones compared to previous web-enabled mobile phones (e.g., “flip” phones; Nielsen, 2011). Still, common web tasks completed on smartphones—via either native applications or the browser—are far less efficient compared to completion times for the same tasks by the same users on the PC (Tossell, Kortum, Shepard, Rahmati & Zhong, 2010). Additionally, page loading delays are problematic to usability (Nielsen, 2011) and lead to sharp declines in usage (Jacko & Sears, 2003).

Finally, smartphones provide almost continuous access to the internet in most personal and professional settings. Without smartphones, PC and laptop users access the internet when they have opportunities to stay at a stationary location for a defined period of time; conversely, smartphone users go online whenever they get the urge (Ericcson, 2011). These interactions occur both on-the-go (Matthews et al., 2009) and at home (Nylander, 2009) usually within a network of other technologies (Kane et al., 2011). Because users carry and access their phones in almost every setting they visit (Ling, 2005), users access the web to fill dead time (Matthews et al., 2009) and perform important tasks that could once only be done on larger computers (Perry, O’Hara, Sellen, Brown & Harper, 2005). Indeed, smartphone users access the mobile web to accomplish specific tasks driven by their context whereas more stationary web use (e.g., on PCs and laptops) is done without regard to use context (van Welie & de Ridder, 2001). As

mentioned in O'Reilly and Battele (2009), smartphones have moved the web from desktops into our pockets.

Taken together, these factors could lead to the development of new patterns of behaviors and routines associated with the use of the internet on smartphones.

## **1.2. Goals**

The goal of this study is to understand the dynamics of these behaviors by characterizing the use of the internet on smartphones and exploring user differences. Because smartphones are used in a wider variety of times and places by a wider variety of users (Hiltunen, Laukka & Luomala, 2002), these devices are indicative of a shift towards the realization of Weiser's (1991) vision of ubiquitous computing. Weiser described a world "activated with technology" to assist users in any environment. His goal, along with those who have followed, was (is) for computers to be completely integrated into all aspects of life. Instead of retreating to a desktop or pulling out a laptop to interact with a computer, users would be surrounded with embedded and networked systems that fit into normal environments. This includes technologies that can be carried on a person across environments. Smartphones provide continuous access to the internet in most environments, yet little research has characterized how the technology is used. In this dissertation, I address three research questions to fill this gap by establishing empirical patterns associated with smartphone internet use and characterizing user differences to inform the design of future systems within the ubiquitous computing paradigm:

1. **How is the web used differently on smartphones compared to the PC?** Previous work in HCI has examined revisitation and search patterns associated with browsing the web on the PC. Very little empirical work, however, has quantified revisitation and search patterns associated with web use on smartphones or where these devices are used. This dissertation will present quantitative results of visits and revisits to internet resources accessed through users' browsers and native applications along with revisits to physical locations. Revisitation and query rates (defined below) have been established for browsing on the PC. In this study, I compare these characterizations with smartphone revisitation rates to understand the recursive nature of smartphone usage and better support user revisiting. Additionally, because smartphones are commonly noted as being operated in diverse locations, I explore the visiting and revisiting of locations where smartphones are used.
2. **How do smartphone owners use both native applications and their browsers to access the internet in real-world environments?** The users in this study are new smartphone owners. In addressing this research question, I examine how these users access the internet through both platforms. Visits and revisits to native applications are examined vis-à-vis their browsers over time. I assess factors such as the nature of the task, experience, user variability and where users access these resources. The results are discussed to understand what each platform affords along with design implications for future smartphone systems and content.
3. **How do users differ in their use of the mobile internet?** Previous research has found that smartphone users differ from each other by orders of magnitude (Falaki,



Mahajan, Kandula, Lymberopoulos, Govindan & Estrin, 2010). However, no empirical research has characterized these differences systematically or examined the influence of experience on this variability. This dissertation examines user variability of internet usage on smartphones (including revisitation patterns, search behaviors, perceived usability and volume of browser use relative to native application use) and characterizes users at two ends of a behavioral continuum, building on previous HCI research. Design recommendations are also suggested for future mobile systems targeting users across the entire spectrum.

The first two questions are aligned with the first goal laid out by Weiser's colleagues (Abowd, Mynatt & Rodder, 2002) to advance his vision of ubiquitous computing: Understand what people do away from their desktops to support daily activities. A consistently large amount of empirical research has focused on how browsers are used on PCs (see, e.g., Weinreich, Obendorf, Herder & Mayer, 2008). Here, I characterize internet use patterns on devices available in almost any setting and the physical localities where smartphones are used. By establishing these behaviors on an empirical foundation, the goal is to better support users in the design of future ubiquitous systems, and particularly smartphones.

The aim of the third question is to explore user variability in these behaviors. A central tenet of HCI research is to understand user differences and design to accommodate users' particular knowledge, skills, age, gender, and income levels among other things (Schneidermann, 2000). In this dissertation, I investigate user differences based on usage of native applications vis-à-vis web browsing to access internet resources along with several other variables. The motivation behind this goal is enhanced design of

future smartphone technology and services to better support users in all personal and professional settings. More theoretically, my goal is to advance HCI research by characterizing different user types in mobile space, building on user differences examined in previous learning and HCI studies.

### 1.3. Approach

A naturalistic and longitudinal approach was applied to address these goals. Real-world use data were logged from iPhones provided to 24 participants to use over the course of one year. The methodology was intentionally designed to decrease participant reactivity. For instance, the logger collected use data unobtrusively and ensured user privacy by obfuscating communications and contact information. These data are first described to holistically understand how users accessed their devices. Additionally, behavior rates are examined and user variance is explored by correlating measures of interest. The results are discussed to understand user differences in mobile space and enhance the design of future smartphone systems and content.

### 1.4. Overview

Before these results are discussed, I review relevant studies of internet use on PCs and smartphones that have informed the present dissertation in **Chapters 2 and 3**. In **Chapter 4**, I describe the naturalistic and longitudinal approach used to collect data along with how the data were preprocessed and measures of interest were calculated. Subsequently, **Chapter 5** presents the usage and survey data collected. Finally, **Chapter 6** summarizes the findings, suggests several design implications, and concludes with

more theoretical implications of this dissertation related to ubiquitous computing and user differences.

## Chapter 2

### USING THE WEB ON PERSONAL COMPUTERS

The roots of the World Wide Web (web) can be traced back to early military research immediately following World War II. Vannevar Bush (1945), a former head of US research and development, recognized that a major component behind the allied victory was the enhanced integration between science and military efforts. He focused scientific resources towards the development of technology that would enhance information access and sharing for effective collaborations. Bush envisioned a time when information could be accessed with a touch of the finger and then shared efficiently with appropriate contacts all from the same device. His proposed technology, the Memex, would allow researchers and operators to access logged information and share that information with others in a network. His seminal work has been attributed to the development of hypertext and the internet (e.g., Buckland, 1992).

Bush's vision of technology-enabled networks of information sharing became a reality with the advent of the PC and subsequent commercialization of the web in the 1990s. The PC allowed users to navigate the internet via a graphical browser. These browsers provided interface mechanisms, such as links and back arrows, which allowed users to traverse web pages identified as uniform resource locators (URLs). Most early and current browsers also allow users to revisit content that they had previously accessed through a history system. These history (or bookmark) systems support users by mitigating three primary web navigation problems (Tauscher & Greenberg, 1997):

First, they can help the user navigate through the vast amounts and poor structure of web

information by providing easy access to information they had visited previously. Second, they can decrease resource use by supplanting search engines for finding old pages, and by eliminating navigation through intermediate pages en-route to the destination. Third, they can reduce a user's cognitive and physical navigation burdens: pages can be returned to with little effort, and they can show users where they have been (pp. 97).

Because of the importance of history systems to support web users on the PC, researchers in HCI and related fields have examined how URLs are visited and revisited to inform the design of these systems.

Indeed, some of the earliest HCI studies of the web used a logs-based approach to this end (Catledge & Pitkow, 1995; Tauscher & Greenberg, 1997). Internet browsers were instrumented with logging technologies to collect naturalistic usage data instead of observing users completing tasks constructed by researchers. From these data, characterizations were developed to establish empirical patterns of web traversing. For example, Tauscher and Greenberg (1997) found that users accessed a small vocabulary of web resources (i.e., unique websites); however, their revisitation frequency within this set was high. They also found that users selected the *back* button often and commonly navigated in a hub-and-spoke pattern. That is, most users navigated to a top-level site and traversed a short distance before returning to this central location and going another direction. Later, Tauscher and Greenberg (1997) analyzed these data and their own web logs by calculating revisitation rates. These logs yielded revisitation rates of 61% and 58% respectively qualifying the web browser on the PC as a recurrent system (Greenberg, 1993) and highlighted the importance of an effectively designed history mechanism to accommodate short- and long-term page revisits.

Because of the changing dynamics of the web, logs-based studies assessing URL revisitation on the PC have been revisited (Obendorf, Weinreich, Herder & Mayer, 2007). Cockburn and McKenzie (2000) collected data for 119 days and found that internet behaviors were even more repetitive than visit patterns found in earlier studies. Users visited an even smaller number of sites than older web studies and reaccessed these sites very frequently. They reported a revisitation rate of 81% from 17 computer science (CS) students. More recently, Obendorf and his colleagues (2007) captured web data from a slightly larger and more diverse set of users ( $N = 25$ , 64% CS). Their logging resulted in a much lower recurrence rate of 46%. They explained that the difference in revisitation rates between these studies and their own study was partly due to the operationalization of how web revisitation rates were calculated. Specifically, they observed that a few older studies truncated the web URL to the base web host to obtain the revisitation measures. This led to substantially higher revisitation rates. Their study obtained revisitation rates considering both the truncated URL (70%) along with the entire URL (46%). Additionally, the reason for the large gap between site (truncated URL) and page (entire URL) revisitation rates, they stated, was due to the changing nature of the web. The dynamic updating of content within sites and larger number of web sites visited were evidence of how the internet has evolved into a more dynamic system. Instead of users frequently revisiting static content, users repeatedly access the same subset of domains to then visit changing content within those sites (i.e., new URLs).

Many activities on the web have also been characterized with Zipf distributions (Adamic & Huberman, 2002). Early studies of the web found enormous variety among users and how sites were visited. Over time, however, patterns were found across studies

showing that there were a few very large elements, some medium-size elements, and a long tail of small elements. For instance, millions of users visit a small set of websites among the billions available (Adamic & Huberman, 2002). A similar Zipf distribution is found for visits to pages within most websites (Glassman, 1994). The most visited pages are inversely proportional to its rank; a few of the most popular pages within the site are visited very frequently followed by a large number of diverse pages visited exponentially less frequently. Zipf distributions have also been applied to caching most popular sites, email networks, and various other internet phenomena (Pitkow, 1998) leading Adamic and Huberman (2002) to state that “on the internet, Zipf’s law appears to be the rule rather than the exception” (pp. 149).

The role of computing experience in web browsing behaviors has also been assessed. Cothey (2002) logged students’ web behaviors over a 10-month period to examine the influence of computer experience. Using a split-half technique, she found that users became more eclectic and passive in their browsing during the second half of the study. Users became more distinct from each other in the web sites they visited, though they generally relied less on search as a function of time. She encouraged researchers to identify characteristics that correlate with behaviors associated with the highly-distinctive web behaviors she reported.

White and Drucker (2007) discovered such characteristics by exploring user differences in searching-related behaviors. They used a logs-based approach to assess intra-user and inter-user variability in search and post-search navigation to characterize these users. Navigators, the anchor name given to users at one end of a behavioral continuum, yielded low intra-user variance; that is, their interaction patterns were

consistent across web use, sequential, and led to high within-session revisitation to web sites. In general, they used the web for concise and targeted interactions. Explorers, in contrast, yielded highly diverse interaction patterns in their searches and post-search navigations. They branched frequently from portal sites, issued more queries in a session and visited more domains in search sessions. Instead of navigating on the web in consistent ways, explorers utilize many different kinds of navigation strategies and largely unpredictable in visiting patterns. Both user types, they suggest, would benefit from different types of interface support. Navigators prefer efficiency to reach a targeted point and would benefit from direct access from a search to the desired information, personal visit history display, and showing portal pages from which these users reach their destinations. Since explorers visited a wider number of domains and valued serendipitous discoveries more, they suggested these users would benefit from guided tours and recommendation systems for displaying potentially relevant links when searches are issued.

Indeed, searching has emerged as a common way for PC users to access resources on the web. Billions of searches are issued each day through engines such as Google, Yahoo, and Bing (Lohr, 2011). A large number of studies have examined query logs to understand the nature of web searching, often in isolation from other navigation sequences (e.g., Jansen, Spink & Saracevic, 2000; Kamvar, Kellar, Patel & Xu, 2009). The results have been applied to develop better query suggestions in search bars (Jones, Rey, Madani & Greiner, 2006), understand the types of information needed when searching (Kamvar et al., 2009), and determine user goals when searching (Rose & Levinson, 2004). Searches from computers are generally short (e.g., averaging less than 3



words per query; Jansen et al., 2000) and done infrequently within a web session (e.g.,  $M = 1.94$  queries per session; Kamvar et al., 2009). Lab studies have found that experts in a domain issued longer searches for information within that domain (Hsieh-Yee, 1993). However, those with more search and domain expertise accessed more websites directly while novices accessed more sites via a search engine (Holscher & Strube, 2000). Those identified as search experts also issued shorter queries. Both logs-based and laboratory studies have shown the effects of user differences, domain expertise, and the type of technology on searching behaviors.

In general, web usage on the PC has been broadly characterized as recurrent (Cockburn & McKenzie, 2000; Tauscher & Greenberg, 1997) and predictable (Pitkow, 1998), but becoming more dynamic, unpredictable, and diverse across users as the web continues to evolve (Obendorf et al., 2007). Web users have been classified as either navigators or explorers based on how they search (White & Drucker, 2007). The former are rare; most users explore the web in ways that are highly serendipitous. This could be due to experience as users become more passive (e.g., issue less searches) and eclectic (e.g., differ from others in visit patterns) as a function of familiarity with the web (Cothey, 2002). The increasingly dynamic nature of the web along with high user variability has led many to suggest that computer users are so unique that individual personalization strategies are optimal for interface support (e.g., Obendorf et al., 2007, Weinreich et al., 2008, White & Drucker, 2007).

## Chapter 3

### SMARTPHONES: CHANGING HOW THE INTERNET IS USED

The studies mentioned in Chapter 2 have provided a thorough understanding of internet use on PCs. However, there are several factors that likely drive differences between internet use on PCs and smartphones. These reasons are discussed here drawing on studies that inform the present dissertation.

#### **3.1. Two Platforms to Access the Internet on Smartphones**

On the PC, the internet is usually accessed through a web browser (Cockburn & McKenzie, 2000). Smartphones provide users two primary ways to access resources on the internet. Similar to the PC, users can get to the internet through a web browser. Many web sites that are opened on smartphone browsers automatically display a version of the site that is optimized for smaller screens, though there are usually options to view the full site if needed (e.g., for functions not available on the mobile version). In addition to this platform, native applications provide access to mobile content designed specifically for the device. These applications are factory-installed or installed by users via an application store (e.g., the Apple AppStore). Once on the smartphone springboard (the smartphone “desktop” display), users can access resources with a touch of a button. The combination of more powerful mobile devices and optimized applications and web content has made using internet resources more efficient in recent years compared to previous-generation web-enabled mobile phones (Matthews et al., 2009; Nielsen, 2011).

Because of this efficiency gain and the optimized design of native applications for each device, recent news stories have declared that “the web is dead” (Anderson & Wolff, 2010) and this has been corroborated with analytics data (Newark-French, 2011). The latter have shown that users spend more time on native applications (internet and non-internet), whether on smartphones or tablets, compared to web browsing on PCs or laptops. They also found that the average smartphone user spends roughly 20 more minutes on native applications compared to their browsers because of the time consumed on the Facebook application.

Though Falaki and his colleagues (2010) did not compare browser use with native application use, they found large variance between users in the use of all mobile applications. They logged native application usage from a number of Windows Mobile and Android smartphone users in the wild. Users differed from each other by several orders of magnitude in the number and type of applications downloaded. For instance, the number of applications used on a given day ranged from 10 to 90 depending on the user. The average amount of time of these interactions varied from 10 to 250 seconds. Systematic differences between users, based on interactions or user demographics, were not found in this research. However, smartphones were used by many to access the internet via both native applications and browsers; outside of communications applications (i.e., SMS, voice phone and email), internet browsing and social networking consumed most usage across both of the datasets studied.

More broadly, users everywhere seem to be choosing to access internet resources through native applications over their browsers. The year 2011 marked the first time ever that the number of smartphone users that accessed native applications surpassed the

number of smartphone users that accessed a mobile web browser according to analytics data (ComScore, 2011). Additionally, both time spent on native applications and the number of native application users is increasing dramatically compared to browser usage (ComScore, 2011).

### **3.2. Mobile Usability**

Still, there are noted usability challenges with the mobile web. Usability, in general, is a term that has evolved within HCI research. Schneiderman has identified five components of usability based on several measures: time to learn, speed of performance, rate of errors by users, retention over time and subjective satisfaction (Schneiderman, 2010). Similarly, Nielsen (1993) defined usability with five concepts of interest: learnability, efficiency, memorability, errors and satisfaction. According to ISO-9241-11 usability is defined by how effective, efficient and satisfying user interactions are with a particular system. Mobile usability has seemed to become a specialty within the more general fields of HCI to assess many of these usability components applied to mobile technologies (Kukulska-Hulme, 2007).

However, mobile usability has been considered as poor for several reasons. First, the lifecycle of many smartphones is short. Before users get familiar with their devices and fully exploit their capabilities, they replace their devices with newer mobile technology (Gilbert, Sangwan & Han Meilan, 2005). Second, challenges with using the web on smartphones result from the size of the device. Text optimized to fit on the small screen interfaces can be difficult to read (Duchnick & Kolers, 1983). Sugden (2005)

reported this was the largest concern of college students' use of a PDA for educational activities. Additionally, text entry limitations are even more problematic for efficient interactions with the web (Tossell et al., 2010). Differences between computing technologies (the smartphone and PC) in efficiency increased as a function of the amount of text entry required for the task. That is, completion times for tasks that required large amounts of text entry were much slower on smartphones compared to the exact same task completed on larger PCs in a laboratory. In common tasks on the web, such as sending an email, searching, and purchasing, users are more efficient on native applications compared to mobile web browsing (Tossell, Kortum, Shepard, Rahmati & Zhong, 2010). A recent study from Nielsen (2009) also found that users were more effective at completing a wide range of tasks through native applications versus browsers on their smartphones. Over 100 users successfully completed more of these tasks on native applications (76%) compared to mobile-optimized sites on their browsers (64%). Based on these standard usability measures of usability, native applications appear to be more usable than mobile web browsing. Both platforms, however, are far less efficient compared to PCs (Tossell et al., 2010).

Third, long page loading times appear to be the primary usability problem with mobile phones (Roto & Oulasvirta, 2005). These delays can result in a sharp decline in amount of usage (Nielsen, 2000) and considered the largest HCI concern in web usability (e.g., Jacko & Sears, 2003). In the 1990s, long delays in web page loading were the most common problems with web use on PCs (Pitkow, Kehoe & Rogers, 1998). Smartphone page loading delays are the primary bottleneck in efficiently using the device (Roto & Oulasvirta, 2005) and lead to much longer non-interaction "wait times" compared to

current PC-based web use (Tossell et al., 2010). According to Wang and his colleagues, these delays are due to round-trip time (RTT), low hardware processing power, and the way resources are loaded after being launched (Wang, Lin, Zhong & Chishtie, 2011). The weakest link in the transmission paths from smartphones to the cloud or server (i.e., RTT) is a primary determinant of page loading delays, independent of the type of internet/data connection speed. While faster mobile network technologies are being developed and deployed (e.g., 4G data), spectrum limitations are likely to keep page loading delays a primary challenge (Hazlett, 2001). More suggestions to improve this usability problem include the use of search to transport users to targeted information without unnecessary navigations through a site (Galletta, Henry, McCoy & Polak, 2006) and predicting usage based on context to preload desired information (e.g., using context-aware systems; see Chen & Kotz, 2000 for a review).

Finally, factors such as brightness (Corlett & Sharples, 2005), weather conditions (Manolo, 2005) and physical movement (Kjeldskov & Stage, 2004) make it even more difficult to interact with smartphones. Taken together, usability associated with mobile web interactions is generally poor (Weiss, 2002). However, the rapid development of new displays, faster ways to input data, predictive systems, optimized interfaces and more powerful processing hardware may have at least slightly mitigated many of these concerns in mobile space. Many of the studies reported above have used devices that may be considered previous-generation devices (e.g., PDAs) and usability concerns may not be as severe or problematic with current-generation devices.

Indeed, users regularly access the internet through their current-generation smartphones. A large number of low and middle income users in urban areas do not even own a larger computer (e.g., PC, laptop, etc.) because of the sufficiency of smartphones to handle their web needs (Fjord, 2011). According to a recent Pew study (2010), one in four smartphone owners (i.e., over 22 million people) prefer accessing the internet through their smartphones over the PC. Many in this group (33%) have discontinued web connectivity at their residences because of smartphone accessibility. Even with connectivity, users still frequently choose to access the internet at home through their smartphones over their PCs (Nylander, Lundquist & Brännström, 2009).

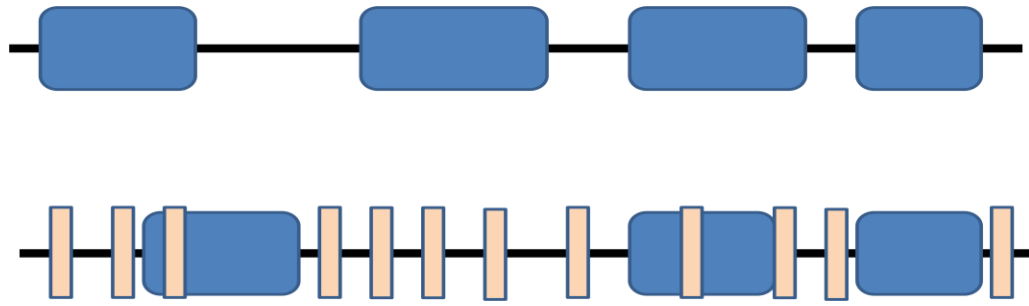
### **3.3. Context and the Nature of Smartphone Interactions**

Smartphones are also widely used on-the-go in diverse environments (Taylor, Samuels & Ramey, 2009). Without smartphones, PC users access the internet when they have opportunities to stay at a stationary location for a defined period of time; conversely, smartphone users go online whenever they get the urge (Ericcson, 2011). Context is important to understanding the nature of these interactions (Chen & Kotz, 2000). Similar to the term “attention” within the field of cognitive psychology, the definition of the term “context” and elements that are included therein are somewhat controversial and vary extensively across studies in HCI (see Tamminen, Oulasvirta, Toiskallio & Kankainen, 2004). In computer science, context can include almost anything that characterizes the situation of some agent (e.g., Dey & Abowd, 1999). This agent can be a person, place or thing. Kim and his colleagues (2002) defined context similarly for

the use of mobile internet resources as “the full set of personal and environmental factors that may influence a person when he or she is using a mobile internet service” (pp. 2).

Smartphones can be accessed in most settings and at any time. Thus, they are used in more diverse use contexts compared to the PC (Figge, 2004; Lee, Kim & Kim, 2005) as the latter is mostly used in predetermined environments (Hiltunen, Laukka & Luomale, 2002). Information needs are not typically driven by context on more stationary computers (e.g., PCs, laptops), whereas mobile internet users accomplish tasks driven by needs in particular environments (van Welie & de Ridder, 2001). For instance, regular smartphone use is developed through habitual routines prompted by contextual cues, such as entering a particular location (Oulasvirta, Rattenbury, Ma & Raita, 2011). In their research, Oulasvirta and his colleagues discovered these habits were often short visits to an application; 18% of Android interactions consisted of users launching a single application, typically either email or Facebook, and then immediately turning the display off. This “checking” behavior was common across users and twice as prevalent for interactions with smartphones compared to laptops. These brief and repetitive interactions were also spread out more consistently throughout the day compared to laptop computing. Similarly, other research (Ericcson, 2011) has shown how smartphone internet access is changing daily patterns of computer usage in a hypothetical day (Figure 1). Without smartphones, computer interactions (indicated by the larger rectangles) are concentrated at several time periods throughout the day. With smartphones, regular and short interactions decrease the amount of time spent on other computers and frequency in which they are accessed. Many times smartphone users access the web in between planned activities (Perry, 2005) or to fill unexpected dead-time (Matthews et al., 2009).





**Figure 1. Hypothetical day of computing without a smartphone (top) and with a smartphone (bottom) based on survey data (Ericcson, 2011). Thinner rectangles represent smartphone internet visits whereas thicker rectangles represent PC usage.**

Context also largely determines the type of mobile service needed by users and the nature of their activity (Lee et al., 2005). Sohn, Li, Griswold and Hollan (2008) used diary studies to discover that users accessed their phones for information because they were cued by something in their environment. The phones used in the study were found inadequate to meet these information needs. Unfortunately, there was no information on the type of phones that their users operated, limiting the interpretability of the results. Regardless, they noted the use of context by the device would be beneficial to address user concerns. They classified the influencing contextual factors into four categories: Activity, Location, Time, and Conversation. Devices that could exploit these types of information (i.e., context-aware computing) were suggested to enhance usability and usefulness of handheld mobile computing.

A growing amount of research in mobile computing has focused on the use of context to predict needs and shift the burden of certain activities to the device and away from the user (Lee, 2010). Just as Google uses previous search history and location to provide more tailored search results and reduce the amount of input required to navigate

to targeted information, context-aware systems propose to leverage contextual information to provide users needed resources within a particular environment. For instance, in their review of these capabilities, Chen and Kotz (2000) detail a number of mobile applications that use location, lighting levels, and user activity to deliver optimized services. More recent work uses environmental information along with information stored on the device (Lee, 2010). Calendar information, for example, automatically detected by the device, can prompt mobile phones to enter particular modes (e.g., Khalil & Connelly, 2005). The word “class” or “classroom” can be detected in an appointment line within the calendar application; using the date and time of the appointment from their calendars, the device is switched to silent mode without any user interaction.

With current smartphones, interruptions both external to the device (e.g., the bus arriving) and internal to the device (e.g., an incoming text message) are regular (Abowd & Mynatt, 2000). Many interruptions do not stop user interactions, but briefly distract users from their devices (Oulasvirta, Tamminen, Roto & Kuorelahti, 2005). Many times the distractions are self-generated. For instance, user attention is shifted away from their devices frequently in mobile contexts for as long as 80% of all page loading durations. Attention was diverted much more as a function of the number of people in the environment. Though users get distracted on PCs as well, interactions with smartphones are typically considered fragmented because of the repeated distractions in mobile contexts, though users adapt and develop strategies to use their devices in these environments (Oulasvirta et al., 2005). One such strategy emerged when they recorded smartphone users walking on a busy street. Users moved their attention (as measured by

eye-tracking) to the environment during page loading and then fixated again on their device once they estimated page loading was completed.

In other ways, the nature of smartphone use is becoming more similar to larger computers. For instance, data from Google search logs (Kamvar et al., 2009) revealed that iPhone users' searching behaviors are more similar to PCs compared to previous-generation mobile phones (e.g., "flip" phones). For example, there were no differences in the average number of words used for search queries between smartphones and PCs. Additionally, the distributions of query topics accessed by their users (entertainment, sports, travel, news, etc.) were similar. They also found several differences between technologies. First, iPhone searching resulted in more diverse information requests across categories per user. Second, computer-based searches resulted in more frequent returns to search results compared to iPhone searches. Third, there were many more queries conducted per session on computers compared to iPhones. According to the authors, this could have been due to information needs that required a quick answer on the iPhone while more in-depth research was saved for computers.

Certainly, individuals use their smartphones to extend and complement their computing on other devices. 75% of the domains visited on users' smartphones were also visited on their PCs (Kane et al., 2011). They suggested use of computer bookmarks or frequently-viewed sites from the latter could be linked to smartphones for enhanced browsing. Bales, Sohn and Setlur (2011) found that iPhone users accessed web browsers more on their phones compared to their computers. The opposite trend was found for Android users. Matthews et al. (2009) outlined several user motivations driving

smartphone behaviors. These include seeking contextually relevant information, completing a concrete task, entertainment, maintaining social ties, and staying aware of potentially changing information. In addition, users optimize tasks for the technology. Many times they optimize what they do on their PCs or laptops as opposed to their smartphones. For instance, users choose to compose longer emails on a PC, but use their smartphones for shorter messages (Matthews et al., 2009). Mobile internet use is generally directed and shorter (e.g., fact finding) rather than the less-directed (e.g., browsing) whereas the reverse has been found for the stationary web (Cui & Roto, 2008). More recent research has shown that this could be a function of user goals in a particular location and not necessarily the technology being used (Lee et al., 2005). Taken together, smartphones are not used “on an island” (Matthews et al., 2009, pp. 1) but within a network of other computers and in a number of different contexts.

### **3.4. Chapter Summary and Look Ahead**

Clearly, Weiser’s vision of digital information at our fingertips to support users in context has become a reality. What started as a vision to support knowledge sharing and secure information transfers in the military has grown into a ubiquitous information portal that can now be accessed by devices that fit into a pocket.

Previous characterizations applied to the PC may not adequately describe how the internet is used on smartphones. Additionally, it is unclear if there are systematic differences in how users interact with the internet via their devices in the real-world. Filling these gaps is the present goal. To this end, I characterize internet use on

smartphones by analyzing naturalistic and longitudinal data collected from iPhone users in the wild over a period of one year. Since there are two general ways to access the internet on smartphones—via native applications and a web browser—I describe usage patterns in each platform along with the influence of experience, the nature of the task and physical locations where smartphones were used on these patterns. Findings are compared with characterizations from empirical studies that examined web use on the PC. User differences are also examined in an exploratory way. The motivation behind these goals is to understand how users access the internet with technologies that are continuously available away from the desktop and inform the design of future smartphone systems.

## Chapter 4

### METHODOLOGY

This study aimed to reveal what users do on their smartphones in real-life environments. Thus, a deliberately naturalistic and longitudinal logs-based approach was applied to collect the data examined in the present dissertation. This chapter describes the methodology in detail. First, I explain how the current methodology was designed to preserve realistic behaviors. Second, I present the methods used to collect data from users' smartphones in the wild including the nature of the logger and how privacy was preserved. Third, I describe the process used to preprocess and analyze the data. Finally, I compare the current approach with several other studies that have used logging technologies.

#### **4.1. Overview**

Smartphones are used in diverse settings (Kjeldskov & Graham, 2003; Ling, 2005) to do a wide array of activities (Falaki et al., 2010). To understand interactive behaviors with these devices, traditional research methods (e.g., laboratory, field, etc.) are commonly used or adapted to fit mobile environments (Hagen, Robertson, Kan & Sadler, 2005). Each of these methods offers several benefits and limitations (Kjeldskov & Stage, 2004). For instance, laboratory studies are highly controlled and can provide data high in internal validity (Rubin, 1994). The potential drawback, however, in many laboratory studies is the lack of ecological validity due to the artificial setting (e.g., Hagen et al., 2005; Wickens & Hollands, 2000).

Because of this drawback, field studies generally provide a more ecologically-valid method when studying mobile phone or smartphones users (Grudin, 1988; Lee, Lee & Kim, 2005). However, field studies have to face challenges when applied to studying communications outside of the lab. Observer effects are a primary concern. Hagen et al. (2005) described several examples when invasiveness adversely impacted the validity of communications data. These field studies required researchers to view confidential meetings, teenagers in their bedroom, and the communication of lovers. Second, traditional data collection techniques have largely required user inputs which may adversely influence accuracy. Diary studies, for example, often interrupt users from their main task, place a burden on participants to report, and rely on their memory of events (Carter & Mankoff, 2005). Finally, because of the mobile nature of smartphone use, shadowing users is very difficult and results in only a fraction of what users actually do on the mobile internet (Palen & Salzman, 2002).

Logging methodologies have addressed many of these concerns by allocating the task of observing to technologies instead of humans. These methodologies provide access to data that can be collected without an observer present or a requirement for users to provide self-reports (Raento et al., 2009). Tasks do not have to be constructed by the experimenter. Instead, data can be pulled from participants' daily activities on familiar interfaces within normal contexts (Kamvar et al., 2009). Thus, data collected from loggers are typically considered more objective, accurate, and realistic (Eagle & Pentland, 2006; Kivi, 2007).

Still, similar to traditional field methodologies, reactivity (i.e., a modification in behavior as a consequence of being measured; Sykes, 1978) can occur if careful steps are not taken to plan for the “selection, provocation, recording and encoding of behaviors and settings” (Weick, 1968, pp. 360) in a way that preserves naturally occurring behaviors. This could seriously impact both internal and external validity of the data collected via these devices (Haynes & Horn, 1982).

For this reason, nine design factors were assessed for the data collection effort (Table 1). These considerations are not necessarily unique to smartphone logging. Indeed, some are shared by other methodologies while others apply more directly to smartphones. For instance, smartphone loggers can collect the content of communications including from people outside of the study (Raento, Oulasvirta, Petit & Toivonen, 2005). Collecting these data, however, can adversely impact user behavior because users may be reluctant to engage in highly personal communications if their privacy is not guaranteed (Purcell & Brady, 1965), even though they may adapt (Johnson & Bolstad, 1975). Capturing communications content may not be considered invasive for dedicated professional systems (Foltz & Martin, 2008); however, communications data on smartphones are highly private in nature (Hakkila & Chatfield, 2005).

Similarly, measurement intrusiveness increases participant reactivity (Sykes, 1978). For instance, novel or changing technologies or interfaces, increased meetings with researchers, and requirements to report can remind participants that they are being studied. This can adversely affect data validity by producing false rates of behaviors and



**Table 1. Nine considerations used in the design of this study.**

<b>Label</b>	<b>Naturalistic Study Design Factors</b>
Privacy	What constraints should be implemented to preserve users' privacy? Many actions performed on smartphones are considered private.
Variables	What variables are needed? Logging is selective regarding what is collected and what is ignored. For instance, researchers can collect a large number of contextual and demographic variables to a small number (e.g., just time and type of searches).
Obtrusiveness	How do I collect the data? This can range from fully-automated (low interruptions) to requiring participants to report (e.g., experience sampling with logger). Survey requests & additional meetings with participants may be considered obtrusive.
Interface	What interface(s) will participants use? New interfaces can be introduced or logging can be embedded and run as a background process on current interfaces.
Tasks	What tasks will participants perform? These tasks can be completely naturalistic (i.e., participant-constructed) or experimenters can construct artificial tasks of interest.
Technology	What technology is used? Logs can be pulled from public files (e.g., search databases) which would allow participants to use familiar technology. On the other end of the spectrum, researchers can provide new instrumented technologies to participants.
Participants	Who are the participants? Subjects may consist of a random population of people that are totally unaware they are being studied to individuals within an academic department or domain of interest (e.g., pilots) that are highly aware of the measurement.
Setting	Where will the study take place? One benefit of smartphone logging is that interaction data can be collected in real environments (instead of a laboratory).
Study Duration	How long to measure usage? This could range from one task of interest to longitudinal measurements over a period of months or years.

unwanted variance (Haynes & Horn, 1982). Users habituate to being measured over time with a stable device (e.g., Johnson & Bolstad, 1973). Constant reminders that the technology is being logged simply reinforce the feeling of being observed, much like a live observer can adversely impact subject behaviors (e.g., Adler & Adler, 1988).

## 4.2. Data collection

The data collection methodology used here was intentionally designed to decrease reactivity. First, multiple privacy constraints were implemented. Second, the approach was longitudinal with no participant requirements to report and no introduction of novel interfaces during the study. Third, in light of previous suggestions to minimize reactivity (e.g., Johnson & Bolstad, 1975), the rationale for the study and anonymization process was described in detail to participants before data collection began and, following this initial meeting, there was limited contacts with participants. I submit these decisions likely allowed the influence of being measured to abate and, thus, the data collected here reflect more realistic behaviors.

### 4.2.1. Participants

The 24 students ( $M = 19.2$  years old) were recruited from a university campus to participate in the study. None of the participants previously owned a smartphone and only 8% were CS students. There were slightly more males in the study ( $n = 14$ ) compared to females. Participants were not paid with cash for their involvement with the study. Instead, they were allowed to use the iPhones for free during the study period and keep their devices upon successful completion of the study.

### 4.2.2. Materials

The devices given to users were iPhone 3GS smartphones (Figure 2). The iPhones ran iOS 3.1.3 for the entire study period of one year as users were told not to update the iOS on their devices. The iPhone comes with several applications pre-installed (Table 2).



**Figure 2. iPhone 3GS displaying the first springboard page with native applications that come standard on the device.**

These include text messaging, phone, mail (email), Safari (web), and the iPod Music Player. Additionally, iPhone users can customize their phones by installing applications from the App Store. There were over 300,000 of these native applications available during the time of this study; some were free while others required a fee.

Along with the device, participants were given headphones, a USB cord, and charging capabilities (i.e., the items that usually come standard with the device). One year of service with unlimited data, texting, and 450 rollover voice minutes was provided

**Table 2. Pre-installed applications for the iPhone 3GS.**

Pre-installed Applications						
Phone	Voice Memos	Photos	Maps	Calendar	iTunes	Notes
SMS	App Store	Camera	Weather	Compass	Settings	iPod
Contacts	Preferences	Safari	Stocks	Calculator	YouTube	Mail

for each participant. There was no requirement to install unique applications or interfaces for the logger to run. Upon successful completion of the one-year study, participants were allowed to keep their iPhones and equipment.

Before we disturbed the iPhones, a custom-designed logger was installed on each device. The logging software (LiveLab) was developed by collaborators on this project to automatically collect data from all user interactions (Shepard, Rahmati, Tossell, Zhong & Kortum, 2010). Participants could use their devices normally and were not interrupted by LiveLab. It ran as a background process automatically after the phone booted. Every night near three o'clock in the morning, the logger connected to a server to upload collected data. These data were compressed and encrypted before transferred and archived to our server. Additional logger components robustly ensured that all data were uploaded to the server. Table 3 shows some of the data captured daily by LiveLab.

**Table 3. Usage data collected by the logger.**

<b>Web</b>	<b>Apps</b>	<b>Locations</b>
URL	Name	Cell ID
Date/Time	Date/Time	Date/Time
Duration	Duration	

Other data collected by LiveLab has been reported elsewhere (e.g., text messages; Tossell, Kortum, Shepard, Barg-Walkow, Rahmati & Zhong, 2012). To see every variable collected by LiveLab, the reader is invited to view Shepard et al. (2010).

LiveLab preserved users' privacy in several ways. First, the participant numbers (user IDs) that were assigned at the beginning of the study were automatically associated

with their data. Any information that may have revealed identifying information was removed before we saw the data. Second, an encrypted tunnel transferred the data, which prevented unauthorized eavesdropping. Third, and most importantly, the actual content and contact information of emails, SMSs, phone calls, and the address book were not collected. Instead, a two part solution to retain research critical data was employed to avoid collecting sensitive data: 1) contact information (i.e., phone numbers, names, and email addresses) were automatically assigned unique alphanumeric codes and 2) text-analysis, performed on the device, extracted relevant information from communication content (e.g., word count, character count, and emoticon usage). Thus, no potentially sensitive information ever left the phone, but we could link important data together for analysis. For example, when a participant sent a text message to his mother and then called his mother, the same code was assigned to the contact for both transactions. The logger also linked many codes assigned to email addresses with the associated SMS and phone codes as well (via the Contacts Application).

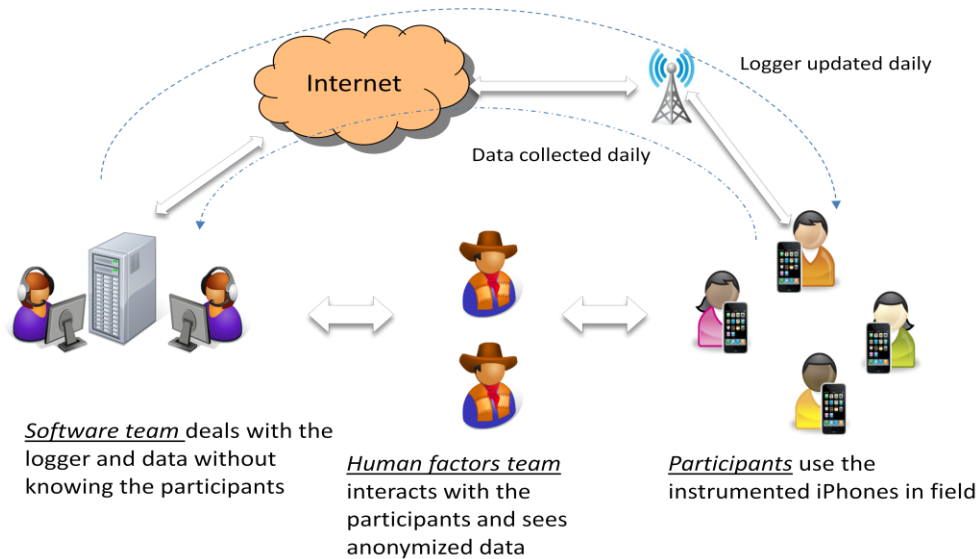
#### **4.2.3. Procedure**

Participants were required to use the outfitted iPhones as their primary mobile phones for the entire year and to simply use the device just as if they had purchased it themselves. Informed consent documents (ICD) outlining the purpose of the study and the type of data collected from participants' iPhones were signed by each user.

Constraints were designed to maintain privacy for times when experimenter-participant interaction was required (e.g., phone malfunctions, questions on service, etc.).

Throughout the one-year study, members of the engineering team were not allowed to

directly interact with users. Instead, the human factors team acted as points of contact for users to interact with for questions or concerns (Figure 3). Both groups avoided associating names with participant numbers. No specific instructions were given about how to use their smartphones.



**Figure 3. Implementation of logging methodology (from Shepard, Rahmati, Tossell, Zhong & Kortum, 2010).**

Surveys were also administered at the beginning, middle, and end of the year-long study. Outside of these three meetings, contact with the participants was avoided. The surveys issued to subjects captured their perceived usability of the device as a whole and helped to interpret logged data. The System Usability Scale (SUS) was used for the former. It consists of 10 questions and has been used widely to assess perceived usability and satisfaction (Bangor, Kortum & Miller, 2008; Brooke, 1996). Other items on the survey required users to perform actions such as send pictures of their bookmarks and springboards to the researchers involved with the study. Many of the survey items were

generated after reviewing some of the logged data to help interpret some of the overall trends.

### **4.3. Data Analysis**

The final logged dataset contained one-year's worth of iPhone usage data for 24 participants. This section describes how the data were analyzed. First, preprocessing of logged data has been noted as a vital step in order to understand user behaviors with the web, though not often reported (Obendorf et al., 2007). Thus, I begin this section with a description of how native application and web logs were preprocessed. The second part of this section describes how the primary measures of interest were calculated.

Descriptive statistics are used to characterize smartphone use along with behavior rates calculated similar to previous HCI studies of PC-based web usage.

#### **4.3.1. Preprocessing the Data**

Since my interest was internet usage, I analyzed only those native applications that required the internet for their primary functionalities. The native applications with functions that largely do not depend on the internet were manually removed from the analysis (Table 4). For instance, the camera application on the iPhone does not require the internet to take pictures or view photos, however it can be used to send pictures over email. Since it is reasonable to assume that cameras are largely used offline, we removed interaction data with this application from before our analysis. I call the internet-connected native applications “native internet applications” (**NIAs**) hereafter. Thus, we

**Table 4. Sample of NIAs analyzed along with several not studied.**

Removed		Kept	
Voice Phone	iPod	App Store	Maps
SMS	Photos	Facebook	Weather
Contacts	Preferences	Games	News

did not consider many iPhone interactions such as voice phone calls, text messages, picture taking, and listening to music on the iPod. However, all new native applications installed by users were considered NIAs for this study such as games and social networking applications. The analysis of NIAs is also handled separately from browser activities for comparative purposes.

Within the web browser logs, all page visits (i.e., full URLs) were recorded. Frequency for page visits is simply obtained by counting the number of visits to URLs. Frequency for site (e.g., domain) visits is counted differently to make comparisons with NIA visits (i.e., launches). Since we did not record page changes within NIAs (e.g., when users went from the News Feed page to Profile page in the Facebook application), we did not count all within-domain site visits in a web browsing session. Adjacent sites visited within a session are counted only once. As an example, Table 5 shows two web browser sessions recorded from one user's iPhone. Sessions are defined as when the browser was launched and then closed. Both the google.com and pensketruckrental.com domains were accessed in each session. The first session yielded two site visits and four page visits and the second session yielded one site visit and four page visits. Similarly, I counted neighboring visits to the same NIA as a single visit as well.



**Table 5. Two web browser sessions recorded from Subject 10.**

1:04:53	<a href="http://www.google.com/search?q=penske...">http://www.google.com/search?q=penske...</a>
1:05:03	<a href="http://www.google.com/aclk?sa=L&amp;amp...">http://www.google.com/aclk?sa=L&amp;amp...</a>
1:05:59	<a href="http://www.pensketruckrental.com/">http://www.pensketruckrental.com/</a>
1:05:51	<a href="http://www.pensketruckrental.com/moving-yourself-truck-rentals/going-green-truck-rentals.html">http://www.pensketruckrental.com/moving-yourself-truck-rentals/going-green-truck-rentals.html</a>
1:06:06	<b><i>End Session – Display Off</i></b>
2:04:55	<a href="http://www.pensketruckrental.com/moving-yourself-truck-rentals/going-green-truck-rentals.html">http://www.pensketruckrental.com/moving-yourself-truck-rentals/going-green-truck-rentals.html</a>
2:05:05	<a href="http://www.pensketruckrental.com/">http://www.pensketruckrental.com/</a>
2:05:34	<a href="http://www.pensketruckrental.com/spsmweb...">http://www.pensketruckrental.com/spsmweb...</a>
2:05:36	<a href="http://www.pensketruckrental.com/Retrieve...">http://www.pensketruckrental.com/Retrieve...</a>
2:05:48	<b><i>End Session – Display Off</i></b>

#### 4.3.2. Primary Measures of Interest

Before I examine into the results, definitions of several quantitative measures are in order. These behavioral rates and indices provide insight into the recurrent nature of smartphone internet use, predictability of NIA use and overall reliance on NIAs relative to browser use. Though descriptive statistics are used heavily to reveal usage patterns from the naturalistic data collected, the measures defined here are also used to establish empirical patterns associated with internet use on smartphones and understand how users differ in their internet visits.

#### 4.3.2.1. Recurrent Behaviors

The recurrence or revisitation rate was originally introduced by Tauscher and Greenberg (1997) to calculate the probability that any URL visit is a revisit. More recently, Obendorf et al (2007) demonstrated the importance of clarifying the terms used in the revisitation rate equation and to report both the full URL revisitation rate and the truncated URL (i.e., site) revisitation rate. I report both of these measures in the next chapter. The page revisitation formula (Equation 1) gives a rate that represents a return visit to the full URL. If any part of the URL is different from the previous visit, it is not considered a revisit. So, for example, visits to <http://www.google.com> and then <http://www.google.com/docs> would not be counted as a revisit using this definition. Instead, the full URL must be reaccessed for it to be computed as a revisit.

---

(1)

The site revisitation rate (Equation 2) is computed similarly. Instead of using the full URL, though, the truncated URL is used. This equation simply captures revisits to domains such as google.com and rice.edu. Similarly, NIA revisitation rates reflect the probability that an NIA launch is a revisit (Equation 3).

---

(2)

(3)

---

Finally, the recurrence of physical location visits is quantified for the first time in this dissertation (Equation 4). The cell IDs that were recorded when users launched their

(4)

---

devices represent locations at a coarse level of granularity. These measures provide a proxy of the recurrence of real-world visiting patterns, perhaps at a similar level as sites on the web relative to pages.

#### 4.3.2.2. *EntroDiversity – Intra-user Variability in iPhone Use*

The EntroDiversity metric is developed to quantify the variability of NIA use for each user, somewhat similar to Kamvar et al.'s (2009) measure to characterize information needs from the web. Whereas the revisitation rates assess how content on the internet is reaccessed, the goal for the current metric, labeled EntroDiversity in this dissertation, is to quantitatively determine variability of how each user employs their iPhone based on their entropy of NIA usage. Highly consistent users (low variability) operate their phones very predictably, using a set of applications in expected patterns. In contrast, highly variable users operate their phones unpredictably accessing a larger set of applications in erratic patterns. In other words, EntroDiversity indices capture whether users operated their phones in more narrow ways by concentrating their application usage

to a small portion of applications or more sporadically across a wider range of applications. The latter would reveal higher variability in how the device is employed.

The properties of the metric (Equation 5) are presented here. I denote  $\{A_1, \dots, A_K\}$  to be the  $K$  possible standard applications and application categories launched by the user. Every user yielded a minimum of 21 for the value of  $K$  as each of the applications that come standard on the phone cannot be uninstalled (recall Table 2). At the time of this study, the Apple AppStore provided the opportunity for users to install native applications from a total of 23 categories (e.g., Games, Social Networking, News, etc.). Thus, an iPhone user in this study that installed an application from each category would yield a total value of 44 for the variable  $K$ . Each user is then associated with a vector  $p_k$  to denote the proportion of the user's tasks in each category. These scores are then be used to develop EntroDiversity indices for each user quantified through information entropy (Shannon, 1948):

---

(5)

EntroDiversity scores return normalized values between 0 and 1 for each participant. A higher value indicates higher variability. For example, if a user opens every application he owned one time each in a given month, this user would be considered the most diverse in their application usage (and yield an EntroDiversity score of 1). In information theoretical terms, there would be little information about this user's needs or preference for an application. The minimum EntroDiversity index (0) would reflect the

reverse. A user that opens up only one application repeatedly and no other applications would be considered the least diverse. In other words, (s)he would be a narrow user and yield the most information regarding needs and preferences for iPhone applications.

For clarity, Table 6 provides a hypothetical example of how these scores are calculated for two different users that visited just two applications over two months (one default application [Mail] and one application category [Games] yielding a value of  $K = 2$ ). User C is the most diverse in terms of application use because she opened both applications at similar frequencies (i.e., showed no preference for one application over the other). This was the case across both months. User A is consistently narrow in that he used SMS in higher concentration compared to AngryBirds. Accessing a larger number of applications in more disperse patterns will increase EntroDiversity scores. Conversely, accessing a smaller number of applications in concentrated patterns will decrease EntroDiversity scores.

**Table 6. Three hypothetical examples of users and their associated EntroDiversity indices. Only two applications are used for clarity alongside (visit counts).**

	NIA Usage		EntroDiversity
User A	Mail (9)	AngryBirds (1)	.47
User B	Mail (8)	AngryBirds (2)	.72
User C	Mail (5)	AngryBirds (5)	1

#### 4.3.2.3. NIA-to-Site Index

The final primary measure used in this dissertation was developed to explore user differences in reliance on browsers and NIAs. Since site and NIA visits were logged at similar levels of granularity, an index was developed using these values for each user

reflecting his or her reliance on each platform relative to the other. The index is defined by Equation 6.

---

(6)

The index values range between -1 and 1. Positive values reflect greater use of NIAs compared to sites on the browser. Negative values show greater use of the browser relative to NIAs. A score of zero reflects that users accessed both NIAs and sites via their browser in the same proportion. This measure is used to discriminate users based on their volume of internet visits through each platform.

#### **4.4. Strengths and Limitations of the Methodology**

Of course, there are several limitations to the above methodology. First, the data from the loggers may not tell the whole story. Even though surveys were used to supplement the loggers, the intent behind the data usage and the perceived user experience of each interaction with the technology were not captured. The immediate context, though we have some variables, is largely missed. Second, not all of the button clicks and display features were recorded. For example, we did not capture how users interacted with the interface to navigate between web pages, such as the frequency with which they selected the back arrow versus a hyperlink. Finally, conclusions from logs-based analyses cannot determine exactly how the interface supported the tasks the users wanted to do at a precise level. For instance, the underlying reasons for users' revisits

were not available from our data. Users could have accidentally pressed the back button resulting in a revisit and logging technologies currently cannot identify these mistakes. While surveys can provide additional clarification on broad aspects of user-interface interaction, approaches such as the other observational techniques employed in Byrne, John, Wehrle, and Crow (1999) can provide better mapping between tasks and interface support.

Despite these limitations, in this dissertation I present the first empirical study of user interactions with the mobile internet using naturalistic and longitudinal data captured from iPhones. Following the research goals outlined in Chapter 1, these interactions are characterized to quantitatively understand how users employ these devices in the wild. The findings are then used to develop HCI suggestions for designing ubiquitous computing systems and understanding user differences.

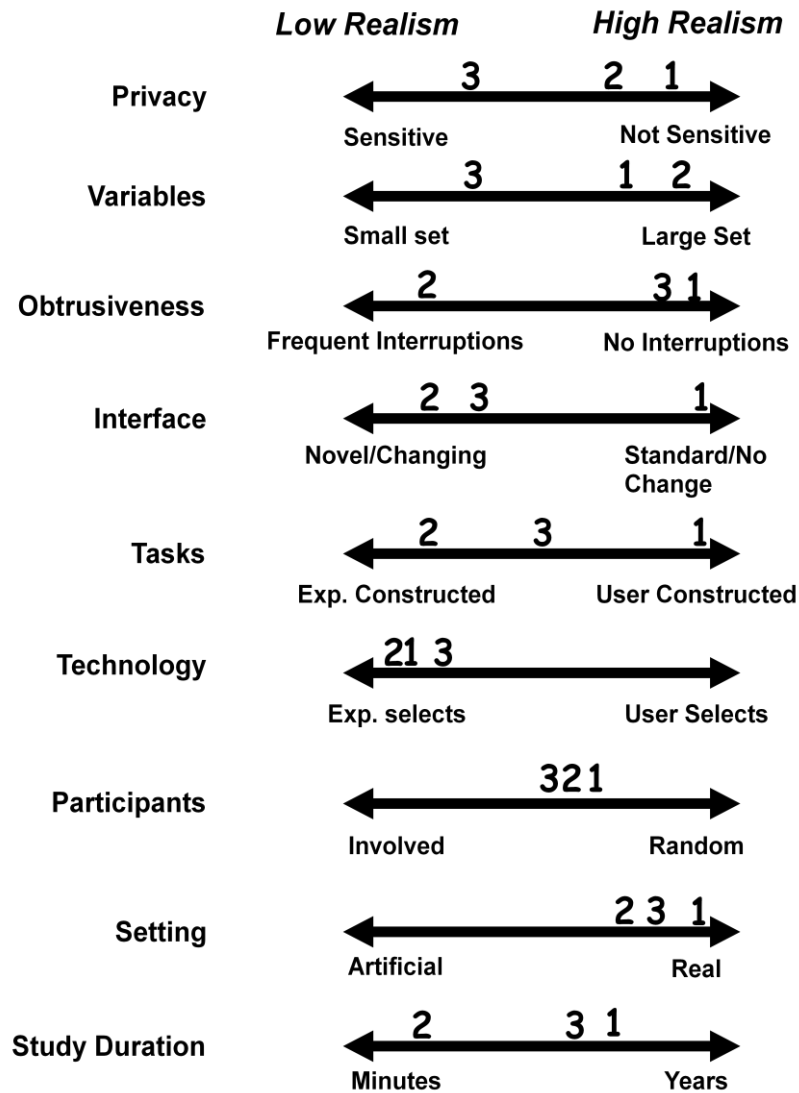
The primary strengths of the approach are the commitment to naturalistic data collection and longitudinal nature of the study. Figure 4 shows where three different studies (including the current one) roughly fall along the dimensions introduced above in Table 1 above. The first study (1) is the current methodology. The second study (2) was conducted by Jönsson, Svensk, Cuartielles, Malmberg and Schlaucher (2002) and used SMS probes to collect information on learning environments for development of distributed pedagogical tools. These probes were sent every day to students' mobile phones at an unpredictable time with instructions. These instructions were in the form of a game or request and had students use their provided phones to collect information (e.g., take a picture of your surroundings). This is similar to other studies that used smartphone

logging for participatory design (e.g., Hulkko, Mattelmaki, Virtanen & Keinonen, 2002). The third study (3) conducted by Oulasvirta, Petit, Raento, and Tiita (2007) used a repeated-measures approach combined with unobtrusive logging to understand communications via smartphones. A combination of experimental control using a standard A-B intervention, a longitudinal collection period (265 days), and the collection of a host of contextual and usage variables truly demonstrates the innovative methods that can be employed with logging (Raento et al., 2009). They also recorded voice phone communications for qualitative analyses. Decisions on each consideration can vary widely across studies, confirming that smartphone logging is a flexible tool that can be leveraged in a number of ways based on research goals. The current approach rates high on most of the dimensions, as seen in Figure 4 suggesting that the behaviors measured were more realistic.

Clearly, this naturalistic approach has provided tremendous access to real-life behaviors with technology in the real-world. Careful steps were taken to implement the “selection, provocation, recording and encoding of behaviors and settings” (Weick, 1968, pp. 360) to preserve naturally occurring behaviors. Quantitative approaches to analyze the data are used to describe what users do on the mobile internet and where they use their devices through a longitudinal lens. Additionally, survey data help with the interpretation of some of the usage patterns collected in the logs.

The aggregate and longitudinal analyses combine with previous research to suggest patterns of interactions and stable user differences associated with mobile internet use.





**Figure 4.** Three studies (1-current study, 2-Jonssen et al., 2002, 3-Oulasvirta et al., 2007) that employed smartphone logging and their approximate placements on each of the nine considerations.

To some extent, the current approach is marked by exploration of broad and realistic patterns of user interactions with technology and not the conclusive elicitation of specific behaviors in controlled environments.

## Chapter 5

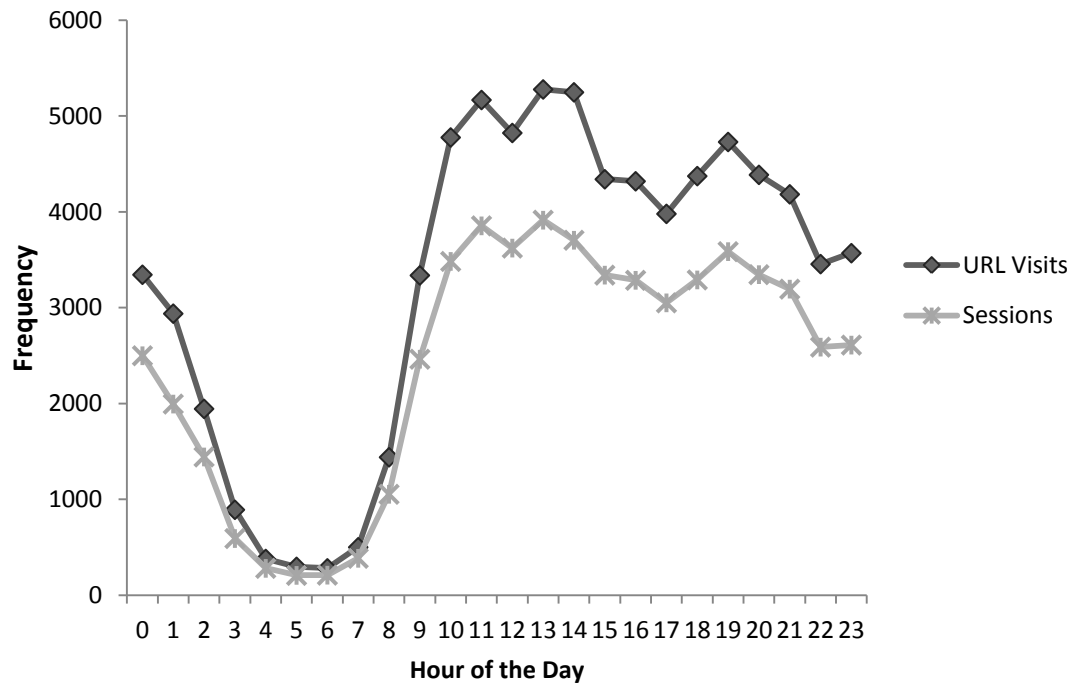
### RESULTS

Over the entire year, NIAs were accessed much more than the browser. A total of 2,080 unique NIAs were launched across 225,151 visits. In contrast, 7,672 URLs were accessed through browsers accumulating 112,083 total visits. This section begins with an examination of visits and revisits to the latter platform (i.e., the web browser) and comparisons with previous PC-based studies. Following this, I characterize NIA and physical location traversing before user differences are explored.

#### 5.1. Web Browser Visitation Patterns

Surprisingly, not all users frequently accessed their browser, even with free service. Users averaged 3.86 browsing sessions per day (*Median* = 3, *SD* = 10.84). The lowest volume user averaged eight browsing sessions per month. He, along with four others, stated a preference for browsing on the computer, low information needs requiring the web, and a heavier reliance on voice phone and SMS. The user that relied on Safari the most launched Safari an average of 11 times per day. Clearly, there was large variance in browsing use among participants.

Most browsing sessions, and subsequent URL visits, occurred during the late morning and afternoon (Figure 5). These sessions were typically short in duration and number of resources visited. The average session within Safari lasted less than two minutes ( $M = 105.86$  sec., *Median* = 96 sec., *SD* = 40.84 sec.) and consisted of a small



**Figure 5. Diurnal patterns of web browser usage.**

number of unique sites ( $M = 2.18$ ,  $Median = 1.5$   $SD = 2.88$ ) and total pages ( $M = 6.07$ ,  $Median = 3$ ,  $SD = 3.58$ ) visited. The differences between the mean and median here reflect the positive skew in the distributions. 52% of web sessions were conducted in isolation from other applications on their phones. Thus, when users accessed their iPhones and visited Safari, it did not include visits to other native applications immediately before or after Safari was launched.

### 5.1.1. Revisiting Pages and Sites on the Mobile Web

The primary PC-based web log studies were completed between 1995 and 2005. By comparison (Table 7), web browsers were used much differently on smartphones. Pages (i.e., full URLs) were not revisited very often (*Page revisitation rate* = 25.3%), yet

sites were revisited frequently (*Site revisitation rate* = 90.3%). The former is substantially less than any of the previously reported study for the PC whereas the latter is *more* than these same studies. On smartphones, just over 60% of all pages were visited once and 15% were visited twice. Interestingly, these numbers are extremely similar to results obtained from the PC 15 years ago (60% and 19% respectively; Tauscher & Greenberg, 1997). Revisitation rate ranges, noting the similar sample sizes across studies in Table 7, were slightly smaller for smartphone users (between 13% and 41%) compared to PC users (e.g., 17% to 61% in Obendorf et al., 2007). The site revisitation rate calculated for browser use on iPhones also yielded low ranges varying between 86% and 97% compared to PC ranges (e.g., 61% to 92% in Cockburn & McKenzie, 2000). Clearly, most smartphone users do not access the same static content on their smartphones compared to PCs, but revisit a set of sites frequently to access a wider variety of pages.

**Table 7. Comparison of primary revisitation studies with the current study.**

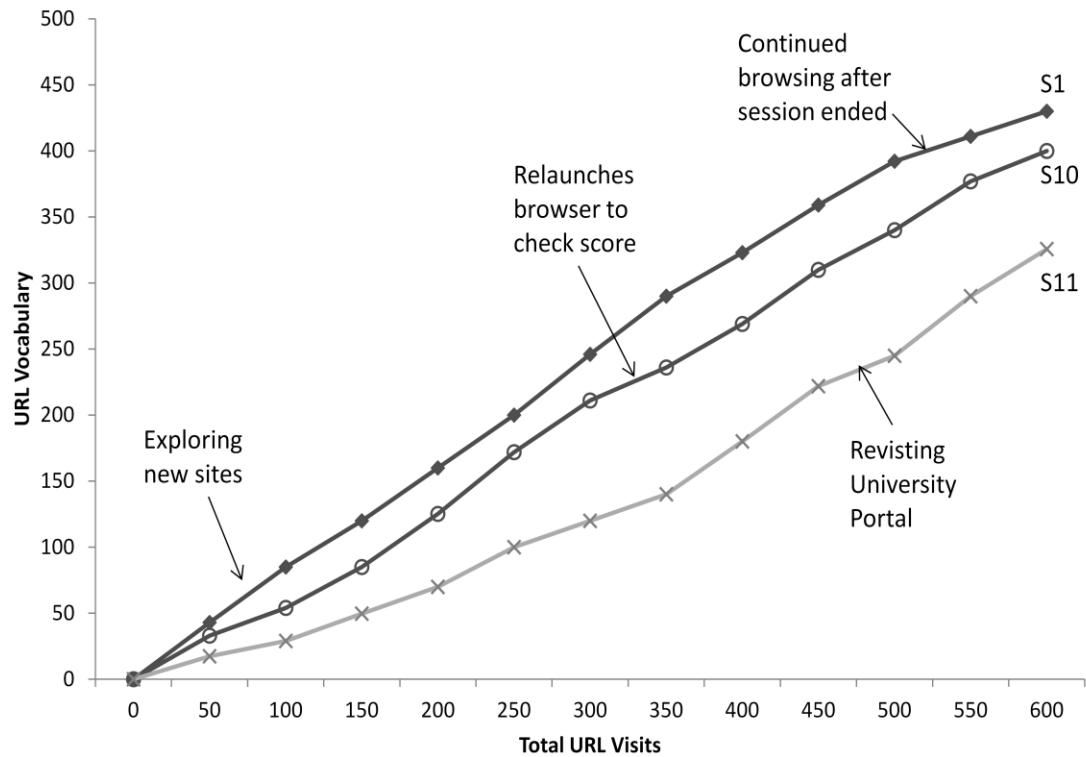
	<b>Tauscher &amp; Greenberg</b>	<b>Cockburn &amp; McKenzie</b>	<b>Weinreich et al.</b>	<b>Current study</b>
Technology	PC - Xmosaic	PC - Netscape	PC - Firefox	iPhone - Safari
Time of Study	1995-1996	1999-2000	2004-2005	2010-2011
Number of users	23	17	25	24
Type of users	100% CS	100% CS	64% CS	13% CS
Length (days)	42	119	105	365
Number of URL visits	19,000	84,841	137,272	79,107
Unique URLs	Unknown	17,242	65,643	58,029
Visits/User	303 - 3299	281 – 23,973	912 – 30,756	156 - 7292
Mean Page Visits/Day	21	41	89.8	9.1 (max = 20)
Site Revisitation Rate	Unknown	81%	70%	90.3%
Page Revisitation Rate	58%	Unknown	45.6%	25.3%

Other aspects of browser usage showed lower diversity among smartphone users compared to the diversity among PC users on the web. For instance, compare the total number of visits logged for both PC users and smartphone users in Table 7. The most recent PC study found an enormous range: From 912 URL visits to 30,756 URL visits. Even though the current study of smartphones lasted three times longer than this study, the range reported here is several orders of magnitude smaller: From 156 URL visits to 7,292 URL visits. The smaller range in the earlier PC-based web studies reflect the small time window in which user actions were recorded. Indeed, the mean rate of URL visits per day shows that the highest-volume smartphone user still visited less URLs per day than the mean rates computed for any of the PC studies.

Smartphone users' unique URL visits relative to their overall visits were substantially higher than PC users. The relationship between the increase of users' URL visit counts and their vocabulary sizes (i.e., number of unique URLs seen) reflects the low probability of a page being accessed again by the same smartphone user. If pages were revisited often, each user's total visit count would grow rapidly as a function of their vocabulary size. This was not the case with smartphones. Using linear regression across all subjects, users' vocabulary sizes strongly predicted total visit count ( $R^2 = .99$ ,  $F(1,22) = 1680.54$ ,  $p < .001$ ,  $b = 1.38$ ). The low value of the slope (1.38) reflects the low page revisitation rate reported above. For every three URLs added to a vocabulary, one page was revisited. New pages were accessed at a higher rate than old pages. There was also large similarity between users in this sample. Slopes from linear regression analyses run for each user revealed a range from 1.01 to 2.79; a smaller range compared to similar

slopes computed for PC web users in the year 2000 (1.87 to 6.46; Cockburn & McKenzie, 2000).

Figure 6 illustrates how three users extended their vocabularies over their first 600 URL visits. Most subjects yielded linear slopes, similar to Subject 1. The variations in the linear slope were slight and typically due to the nature of the task as described more in the next section. For instance, Subject 10 relaunched Safari consecutively (i.e., in isolation of any other application launches) over a short period of time to reload the ESPN.com home page; likely to check sports scores. Fragmented browsing across interruptions also resulted in page revisiting. Many short-term page revisits (accessing the same URL within a period of three days) occurred via a page being reaccessed across adjacent web browsing sessions. I found that 21% of all page revisits were due to the loading of a page previously closed with the browser. Our logger did not record the page visit if it did not load completely. Thus, many users retrieved this information to continue a previous navigation sequence, such as Subject 1 in Figure 6 (also recall Table 5). Roughly one quarter of these revisits were continued after an interruption such as a text message or voice phone call. The other three quarters of these type of revisits were after the phone display was turned off for some period of time. A few users reported using small periods of free time to access a page for viewing later, perhaps a strategy developed to deal with long page loading times. Subject 11 seemed most similar to PC users and yielded the highest revisitation rate (41%). He visited a large number of pages not optimized for mobile browsing. For instance, the Rice University academic portal (i.e.,



**Figure 6. URL vocabulary by total URL visits over the first 600 visits.**

OwlSpace) was frequently revisited by this user. This site requires users to visit the same page to log-in and interact with content. Additionally, the follow-on home page allows users to access content for various classes. Subject 11 visited this page frequently over the entire length of the study.

Though *pages* had a low likelihood of being revisited, the *sites* previously visited by our users had a high probability of being revisited. The site revisitation rate of 90.3% is substantially greater than the page revisitation rate and corroborates previous research suggesting the importance of clarifying what is being measured. This seems especially important for the mobile web. Linear regression across subjects showed that for every

two new sites added to the overall vocabulary, roughly 25 visits were to a site previously accessed ( $R^2 = .51$ ,  $F(1,22) = 27.01$ ,  $p < .001$ ,  $b = 12.60$ ). This set of sites was generally small, though varied greatly between users; the top five sites accounted for a minimum of 22.3% of page visits to a maximum of 91.8% of all page visits with a mean of 73.4%.

#### *5.1.1.1. The Influence of the Task on Revisiting Browser Content*

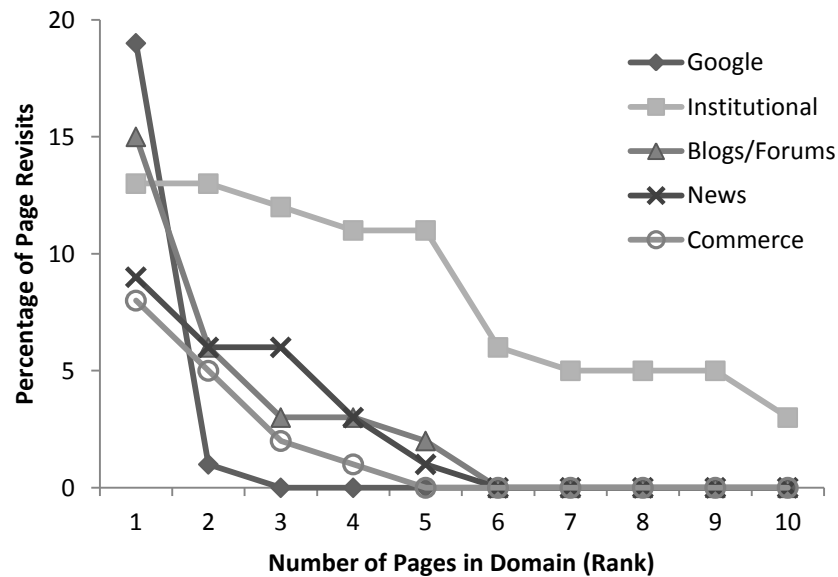
Because of this distinction between site and page revisitation rates, the type of content revisited on the mobile web and how that influenced said distinction was assessed. The URLs visited by our users were classified into eight categories (Table 8). In order to increase reliability in this process, two independent coders were recruited to classify the over 112,000 URL visits. These two coders' strongly agreed with their categorizing based on a Kappa score of .79 (Landis & Koch, 1977). The relatively small number of disagreements were reconciled by the author. Most visits were to five types

**Table 8. Frequency and percentage of URL Visits by Type.**

<b>Category</b>	<b>URL Visits</b>	<b>%</b>
Search	34745	31
Institutional	26899	24
Social/Blog	12329	11
News/Sports	11219	10
Commerce	11127	9
Adult	4922	4
Religion	4483	4
Games/Movies	3362	3
Health	2241	3
Travel	1120	1



of sites. The top five types of sites visited were domains users accessed to issue queries (e.g., google.com), blog (e.g. neoseeker.com), access the News (e.g., nytimes.com), shop (e.g., craigslist.com) and interact with the institutional site for Rice University (i.e., rice.edu). Based on the nature of these types of sites, it appeared many of the frequently accessed sites led users to unique pages. Figure 7 shows that this was indeed the case, though the type of site did impact the number of unique pages visited. For example, the majority of our users (79%) had niches of three highly revisited sites that contained Google.com and Rice.edu. The *site* revisitation rate for Google was high because all search results pages were within the same domain. However, most search results *pages* were unique because of different query terms (and thus unique URLs). Similarly, blogs were revisited often across users and yielded a long tail (Figure 7). In contrast, there were

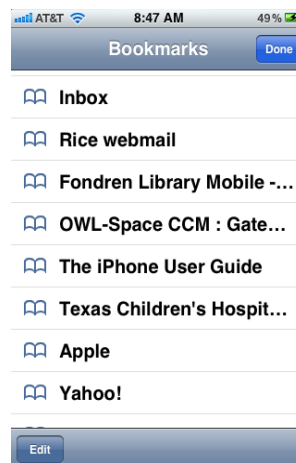


**Figure 7. Distribution of page revisits within domains for each type of site (after Obendorf et al., 2007).**

more sub-top level pages revisited within institutional sites such as the Rice.edu domain (e.g., registrar.rice.edu, dining.rice.edu) and News sites yielding shorter tails.

Since only 9% of all pages were revisited more than five times, the same two undergraduate researchers that collaborated on this project manually categorized these pages as (1) log-in pages, (2) subsequent home pages after a log-in, (3) a top-level page (e.g., <http://google.com>, <http://www.espn.go.com>), or (4) other. Many were indeed in the first three categories (62%) demonstrating that top-level sites were used as a gateway to other content, often within the same site (corroborated by the high site revisitation rates and Figure 7).

Bookmarks within the web browser (Figure 8) were not used frequently on the mobile web. A substantial 83% of users did not add any bookmarks at all in Safari and two of these users did not realize there was a bookmark system within their browser. Three other users only added one or two bookmarks throughout the entire year, but

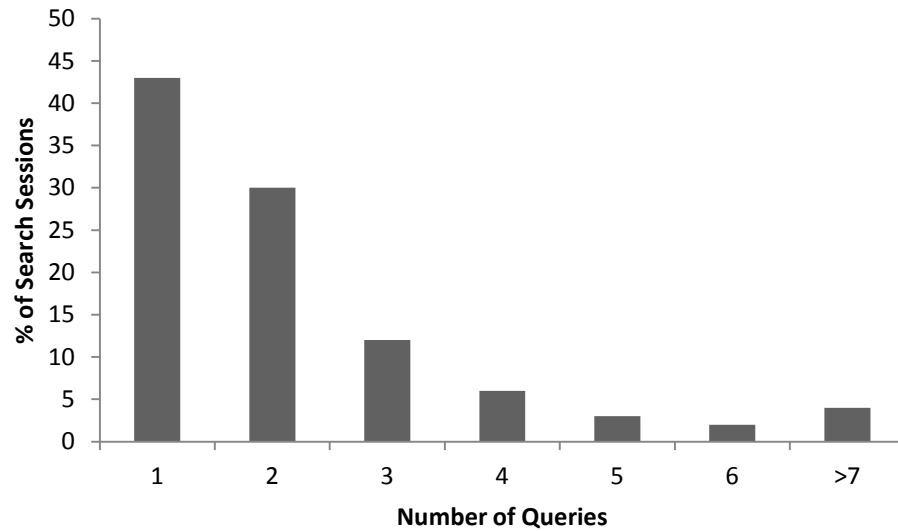


**Figure 8.** Screen capture of the bookmarks window on the Safari browser.

reported they rarely use it and added them immediately after receiving their iPhones. Eventually, these two users added bookmarks to their springboards (one linked a single page while the other linked two pages). No other users bookmarked a page by adding an icon (i.e., link) on their springboard. Based on these data, bookmarks within browsers or adding icons linking content within browsers to springboards did not constitute primary methods for revisiting content on the mobile web. However, 54% of the sites visited also were installed as an NIA (e.g., Google Mobile, Rice, etc.) which appeared to supplant other types of bookmarking to some degree as discussed further below.

#### *5.1.1.2. Searching on the Mobile Web*

Search was also used frequently to visit and revisit content on the mobile web. Users issued over 17,500 searches across the entire study period resulting in a 59% query rate (i.e., number of browsing sessions that consisted of at least one search). Users did not vary much in their volume of searches ( $M = 56.3\%$ ,  $SD = 7.2\%$ ) and use of Google. Less than 0.1% of all queries were conducted outside of Google.com. Most browsing sessions with search (85%) contained less than four queries ( $M = 2.14$ ,  $Median = 2$ ,  $SD = 1.32$ , see Figure 9). As Table 9 shows, differences between technologies were large. Smartphone users issued searches within nearly five times more browsing sessions compared to PC users. However, the number of searches issued within these sessions was half the amount that PC users yielded five years ago. According to subjects, there were two reasons for the relatively low number of queries per search session compared to PCs. The first was due to low time available for extended web browsing. The second was problems with long page loading times.



**Figure 9. Percentage of search sessions by the number of queries issued.**

Google searches were done via the browser bar located at the top of the screen (62%) more than the search bar on the Google site (38%). Browser logs also showed that 39% of the multi-query sessions (i.e., all sessions that included more than one search) were consecutive and contained at least one word from the previous search by the same user. However, our manual categorization process also revealed that many of these (26%) of these were due to fixing mistakes. Additionally, we found a very low percentage of visits to the second page of search results (1.3%). The other 61% of the multi-query sessions consisted of queries for related or unrelated topics that were not consecutive.

**Table 9. Comparisons between PC and smartphone searching behaviors.**

	White & Drucker (2007)	Current Study
Page visits/search session	17.7	7.4
Mean Query rate	12.5%	56.3%
Queries/session	4.3	2.1

Recall that effective history systems can help web users get to targeted information efficiently without the use of search (see quote in Chapter 2 from Tauscher & Greenberg, 1997). Surprisingly, mobile web users often issued searches to return to a previously accessed page. A considerable 47% of all page revisits after three days occurred through users searching and then navigating to targeted content. Indeed, all users but one accessed sites in their top three (e.g., rice.edu, neoseeker.com, craigslist.org) via Google consistently throughout the entire study. Additionally, we asked users, via an open-ended question, to report some of their common strategies for revisiting web pages that they have not been to in two weeks. Every user mentioned the use of search. Interestingly, typing in an address within the URL address bar was mentioned by only one user.

## **5.2. Native Internet Application Visitation Patterns**

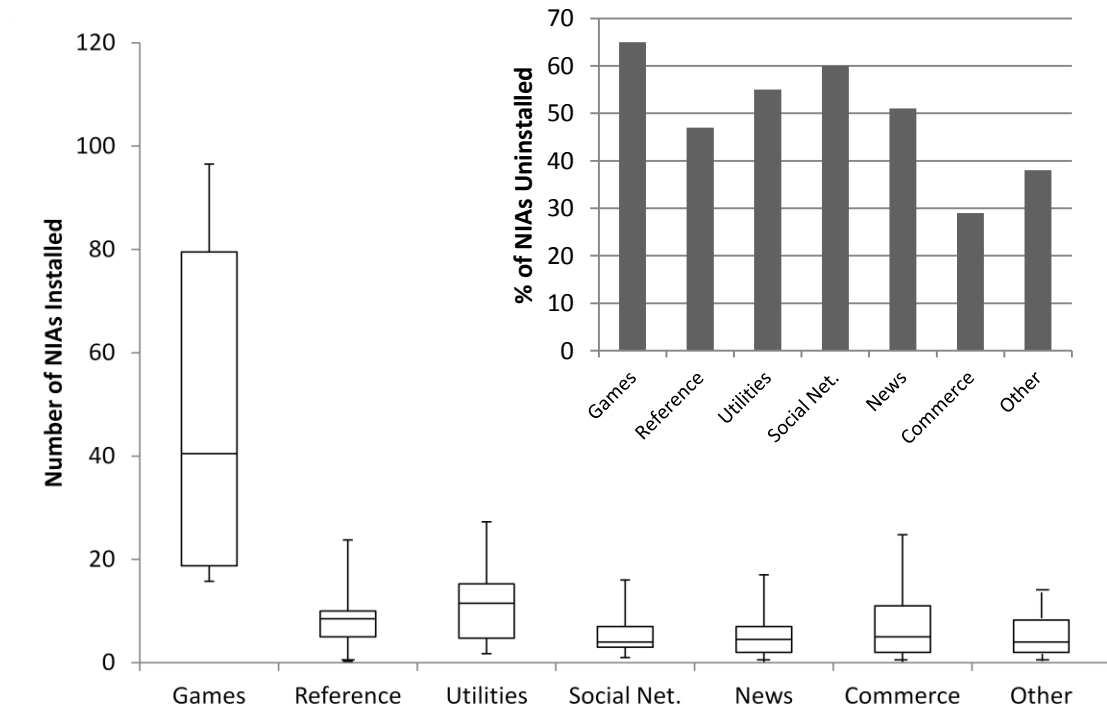
Though NIAs were used more heavily overall, there was still large diversity among users. Table 10 shows aggregate statistics similar to those reported for the browser. Even though vocabularies across our participants ranged from 31 to 475, all of the NIA revisitation rates were above 95% showing the highly repetitive nature of NIA use. The overall revisitation rate of 97.1% was driven by a high number of visits to a relatively small set of NIAs within a user's vocabulary. Also similar to site visits, NIAs were used for brief periods of time ( $M = 137.29$  sec.,  $Median = 62$  sec.,  $SD = 207.29$  sec.). As can be seen by the large difference between the mean and median, this distribution is positively skewed with most NIAs being used for roughly one minute when accessed. Of course, the large standard deviation also reflects the fact that durations

of NIA use are heavily task-dependent and largely a function of what NIAs were installed by each user. For instance, NIAs such as Pandora were used for longer periods of time compared to others (e.g., Weather). Across all NIAs, many were operated independently; 59% of iPhone activations (i.e., the display was turned on) were for the use of one NIA alone before the display was turned off.

**Table 10. Summary and variance statistics for native internet application use.**

	<b>Mean</b>	<b>Med</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>
Visits/Day	37.40	38.01	8.30	1.30	65.89
Vocabulary	124.63	102	106.17	31.00	475.00
Revisitation Rate	0.97	0.96	0.02	0.91	0.99
Hours/Active Day	2.02	1.5	1.21	0.26	4.69
NIAs Visited Once	33.1	28	41.43	0	185
% Search App Use	0.002	0.002	0.004	0.00	0.02
% Visits in Top 10	60.39	67.13	13.15	32.39	89.84

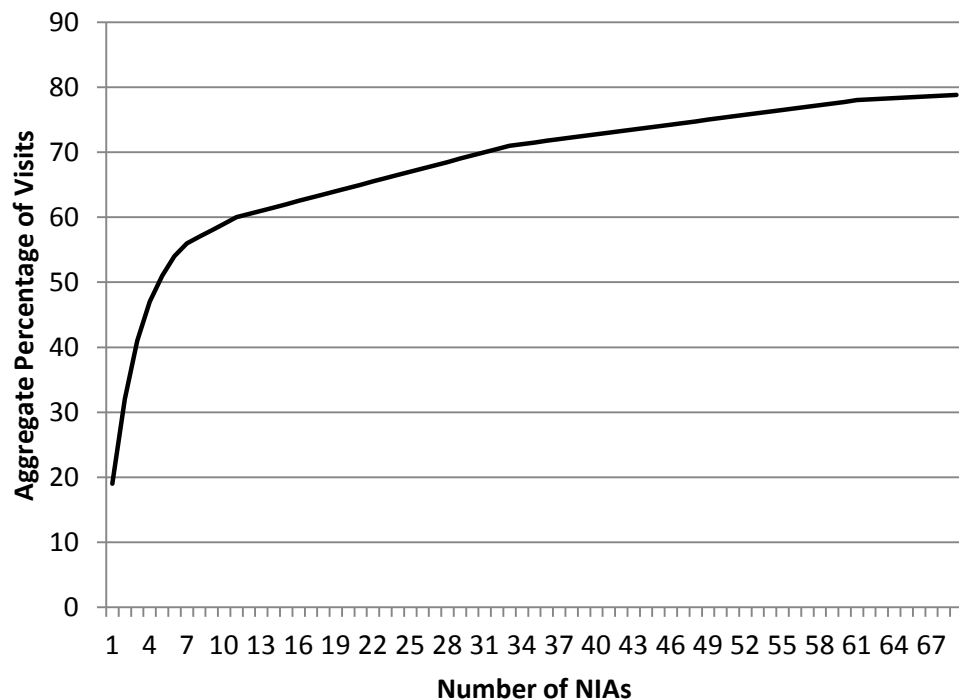
Participants extended their NIA vocabularies by installing new NIAs via the Apple AppStore. At the low end of the spectrum, one user added only 10 NIAs to his smartphone over the course of the 12-month study. Conversely, on the high end of the spectrum, another user installed over 451 native applications within this same duration. Using the AppStore categories defined by Apple, users installed many more Games compared to other NIAs (Figure 10). NIAs in other categories were installed at levels similar to each other. Dissimilar to sites visited on the browser, NIA vocabularies did not



**Figure 10. Number of native applications installed by genre. The inset plots the percentage of NIA uninstalled by the same genre distinctions.**

explain much of the variance when regressed on NIA visits, though the predictor significantly influenced the outcome variable ( $R^2 = .19$ ,  $F(1,22) = 5.18$ ,  $p = .03$ ,  $b = 4.48$ ). Thus, for every seven visits, two new NIAs were installed but this may not apply broadly. Because much of the variance was not explained through linear regression, other curves were applied to fit these data. A cubic curve explained the variance ( $R^2 = .44$ ) better than both logarithmic ( $R^2 = .34$ ) and quadratic fits ( $R^2 = .38$ ), though much variance is still left unexplained. Logarithmic beta weights yielded lower  $p$ -values compared to the other fits ( $p = .02$ ) suggesting that users' recurrence to NIAs was low during initial visits but revisitation rates increased with experience.

As shown in the inset of Figure 10, there was low probability for NIAs to stick on users springboards. Most NIAs added by users were also uninstalled at some point during the study period, though with large user variance ( $M = 82.25$ ,  $Median = 43.5$ ,  $SD = 29.85$ ,  $Min = 0$ ,  $Max = 182$ ). Just because users kept NIAs on their springboards did not mean they were revisited either. Many NIAs, in fact, were visited either once or twice (30%) and then deserted on springboards. Similar to site visits, a handful of top NIAs consumed the majority of all visits (Figure 11). The Mail and Facebook applications were within every users' top 3 NIAs used, though one subject did not install the latter. Other top NIAs included Maps, Words with Friends, YouTube, and the AppStore. NIAs that could be



**Figure 11. Cumulative percentage of native internet application visits.**

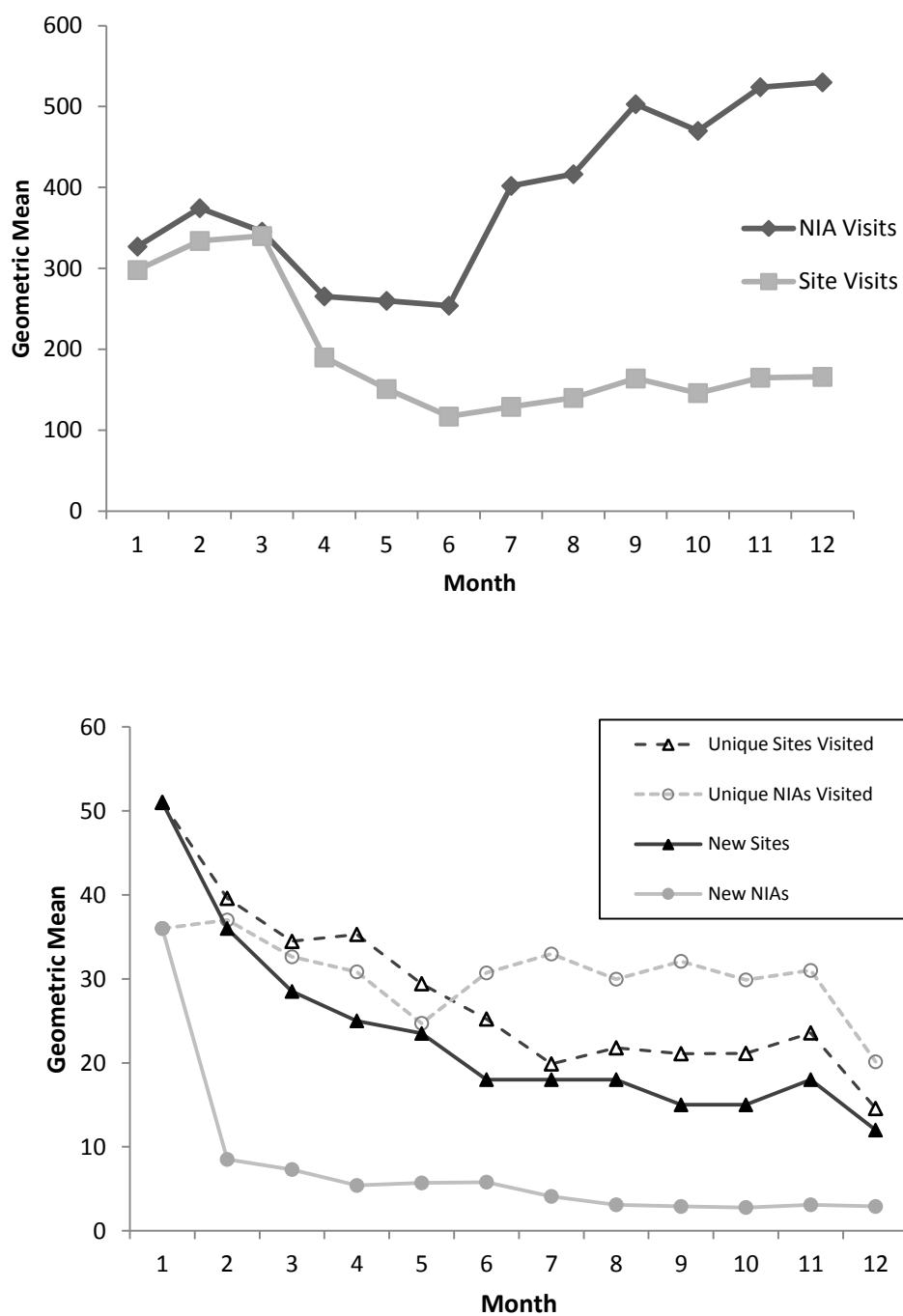


used to issue queries were installed by 91% of our users. Google Mobile was most common, though one user did install Bing as well. Interestingly, these applications were not used very often; across the entire length of the study, they were accessed an average of 6.2 times over the entire year (*Median* = 5.5, *SD* = 4.97).

### **5.3. NIA and Site Visits: Temporal and Contextual Patterns**

From an aggregate perspective, NIAs were visited more than twice the amount of sites. Longitudinally, however, the distinctions between NIA and site visits appeared to be driven by experience and seasonal influence. Figure 12a shows sites and NIAs were accessed at similar levels for the first three months of the study. Months 4-6 were during the summer break. Within this period, all internet use decreased; browsing decreased more dramatically compared to NIAs. After the summer, browser use increased modestly compared to the use of NIAs which sharply increased. Users generally reported the low summer use was due to decreased school-related activity along with being closer to friends and family they usually corresponded with online. Once they returned for the new academic semester (Month 7), the number of visits to NIAs increased substantially while browser use remained relatively low. The large reliance on NIAs continued to grow while browser use was stable across the rest of the study period.

Most of the standard NIAs that come with the device as well as new NIAs from the App Store were explored during users' first month with their devices. New NIAs



**Figure 12. Number of (a) total visits, (b) unique content and new content accessed through NIAs and the browser by month. Geometric means are used in these figures to reduce the influence of outliers within each month on the summary statistics.**

not installed and accessed as frequently as sites. New content was visited more than twice as much on sites ( $M = 27.33$ ,  $Median = 25$ ,  $SD = 10.94$ ) compared to NIAs ( $M = 10.04$ ,  $Median = 9.5$ ,  $SD = 9.69$ ). Instead, a more steady set of NIAs were accessed regularly by our users. The installation of new NIAs dropped more sharply after the first month. By the third month, users averaged under eight visits to new NIAs per month. Figure 12b shows how the distinctions between total vocabulary and new content for each platform changed as a function of experience. The total number of NIAs accessed accumulated with time and by six months NIA vocabularies exceeded site vocabularies. Browsers were continually used to access new content and these new sites generally made up most of users' overall vocabularies within a given month. Users generally toured a large number of diverse sites early on with their device before settling on a handful of top sites that were frequently revisited. The sites that were revisited were revisited very frequently resulting in 3-5 sites that were generally revisited heavily for each user and a long tail of unique sites not revisited after one or two visits. The handful of sites settled on by our users was diverse. While most users accessed the Google and Rice sites regularly throughout the study, the other most dominant sites were unique across participants. Indeed, no pair of users accessed the same site in their top 5 with the exception of Google and Rice.

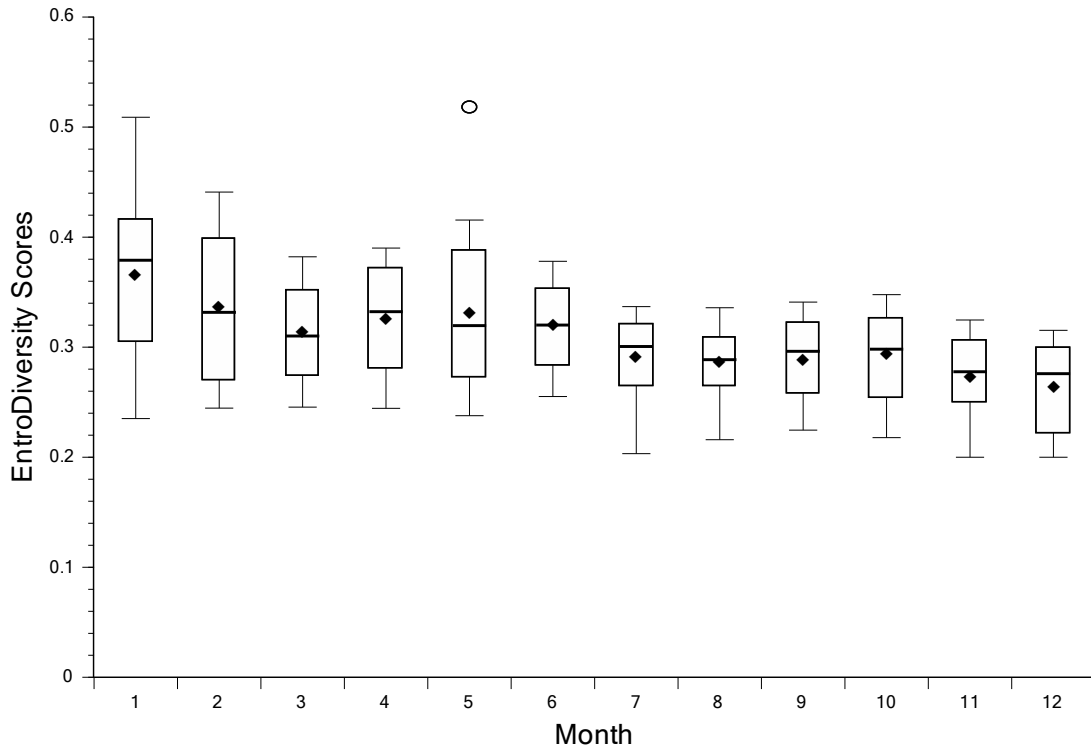
In an open-ended question answered after data logging concluded, users described why they used their browser more to access new domains compared to NIAs. 79% stated they did not want to take the time to install an NIA that they perceived would not be used again. Most of these users reported that they used their browser to get quick information and were unable to predict their information needs to install an appropriate NIA ahead of

time. One user stated this clearly, “sometimes Google knows what I need better than I do...actually, most of the time.” Five users mentioned that installing NIAs was more like shopping whereas accessing new content on the web browser was more obligatory for information needs.

### **5.3.1. Behavior Rates and Indices**

Overall amounts of usage measured by raw data yielded high variance between users (recall Table 10). In this section, I characterize both NIA and browser use by examining the behavior rates and indices introduced in Chapter 3. These measures are described longitudinally by month similar to the previous section. Additionally, analogous to logs-based studies that examined the influence of experience on PC web behaviors (e.g., Cothey, 2002), a split-half technique is used to assess differences between earlier and later behaviors. Scores computed for the first six months of the study are subtracted from scores for the second half of the study. These difference scores are compared to zero using one-sample *t*-tests to assess change patterns associated with the users’ increased experience.

First, intra-user variance is assessed via NIA usage patterns. An individual’s level of entropy is used to this end as measured by the EntroDiversity index defined above (scores range from 0 to 1 with higher scores representing higher within-user variance and lower scores representing more predictable patterns and lower within-user variance). Across the entire period of the study, participants seemed to use NIAs in narrow and predictable ways ( $M = .34$ , *Median*, = .36,  $SD = .06$ ). Longitudinally, EntroDiversity scores decreased and converged over time (Figure 13). During the first month, users’



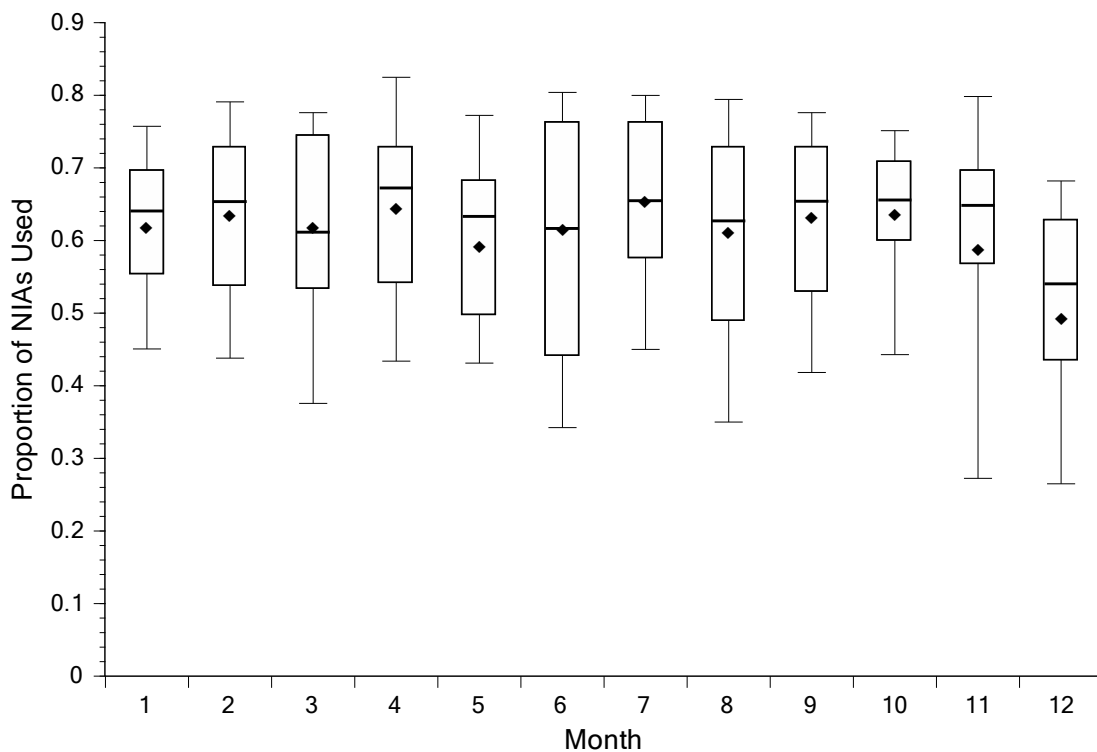
**Figure 13. Box plot of EntroDiversity scores by month.**

EntroDiversity indices were largest compared to other months ( $M = .33$ ,  $SD = .91$ ).

However, levels of EntroDiversity decreased with time as seen in Figure 13. By the final month, both the level and variance of entropy was lowest ( $M = .26$ ,  $SD = .44$ ). Difference scores, obtained by subtracting EntroDiversity scores from the first half of the study from the second half of the study similar to above, showed that users became more predictable in the latter ( $t(23) = 4.00$ ,  $p = .001$ ). While browsing gradually became associated with accessing new content through search, NIA patterns became more systematic and concentrated.

This finding was corroborated by exploring the proportion of NIAs accessed on users' phones. Through the web browser, users can visit an enormous number of URLs. Conversely, the number of NIAs on a user's phone is fixed unless more are installed

through the AppStore. The ratio of the number of NIAs used divided by the total number of NIAs on each user's device showed that users frequently left many NIAs on their devices untouched throughout the study period (Figure 14). Differences between the first

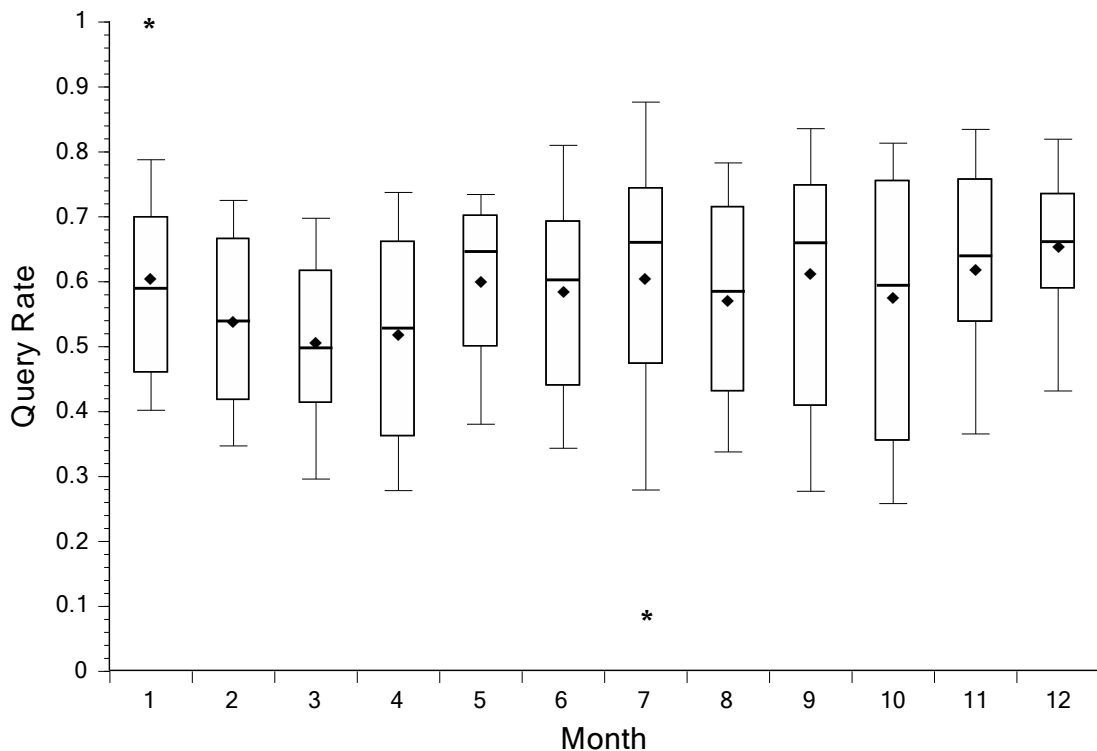


**Figure 14. Box plot of the proportion of NIAs used to total NIAs within each user's vocabulary by month.**

and second halves of the study were not significant ( $t(23) = 1.79, p = .19$ ), perhaps due to the high number of uninstalls mentioned above keeping the denominator of this ratio low.

Search within the web browser was regularly used to access new sites, even though most users installed an NIA to search (e.g., Google Mobile, Bing, etc.). Though

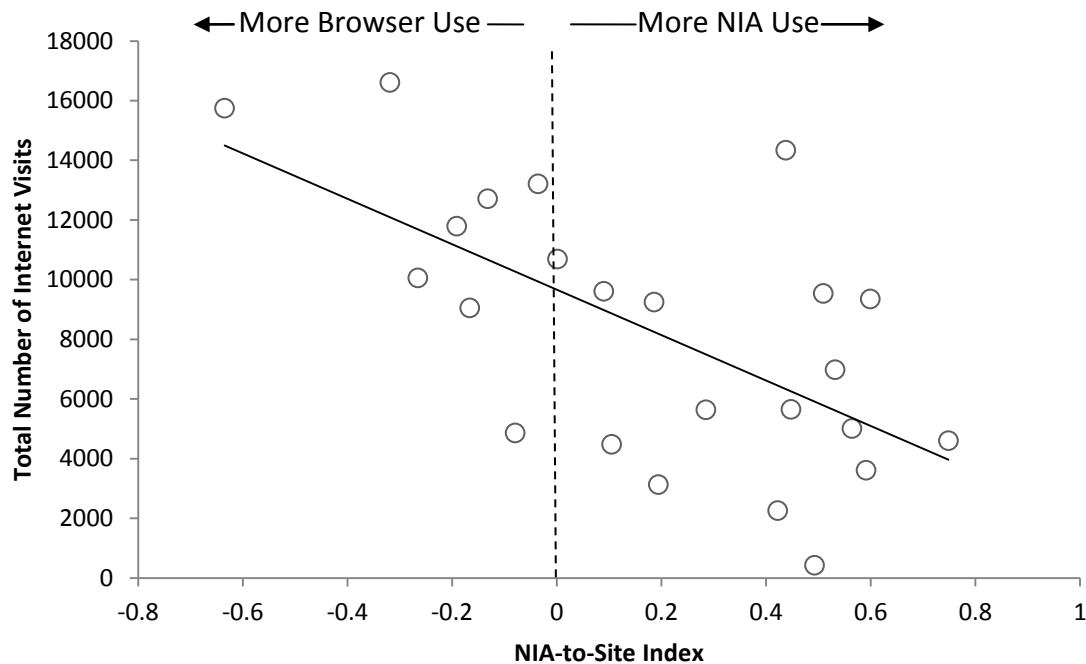
overall visits to sites decreased with time (Figure 12a), use of search, in contrast, remained stable and became proportionately more central to browsing (Figure 15). The number of web sessions with at least one query (i.e., the query rate) increased 5% in the second half of the study when compared to the first half. This increase was significant ( $t(23) = 2.17, p = .04$ ).



**Figure 15. The percentage of browsing sessions that included Google over time.**

Finally, I explored how experience with the iPhone influenced use of the browser relative to NIAs. Surprisingly, not all users yielded higher visit counts to NIAs relative to their visits to sites on their browsers as measured by NIA-to-site indices. As Figure 16 shows, users varied greatly in their reliance on each platform. An inverse correlation was

also found between overall internet use and NIA-to-site indices ( $r = -.61, p < .001$ ). Thus, most participants visited NIAs; however those that used the internet more via their smartphone also more frequently accessed sites via their browsers. I considered that one reason NIA-to-site indices were so low was due to high early browser use. Recall that all



**Figure 16. Scatter plot of total internet visits by NIA-to-site indices**

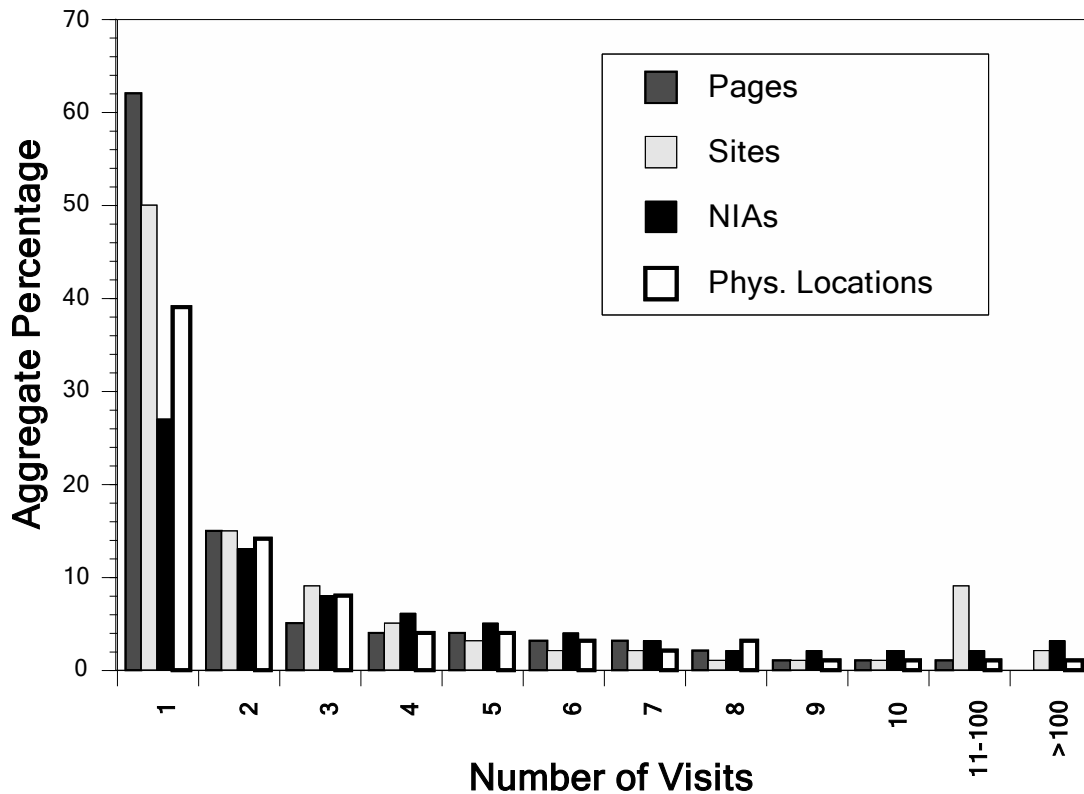
of our users were not previous smartphone owners. Computer experience could have transferred to users' smartphone use and led to higher web browsing early in the study. Thus, differences in NIA-to-site indices were assessed between the first and second halves of the study. Findings show that 20 of the 24 users increased in NIA use over browser use in the second half of the study. The mean of these difference scores ( $M =$



1.56,  $SD = 2.43$ ) was significantly greater than zero ( $t(23) = 3.14, p = .01$ ). Thus, users' reliance on NIAs relative to their browsers increased with experience; however, this was not the case for every user.

### 5.3.2. Physical and Virtual (Internet) Revisitation Patterns

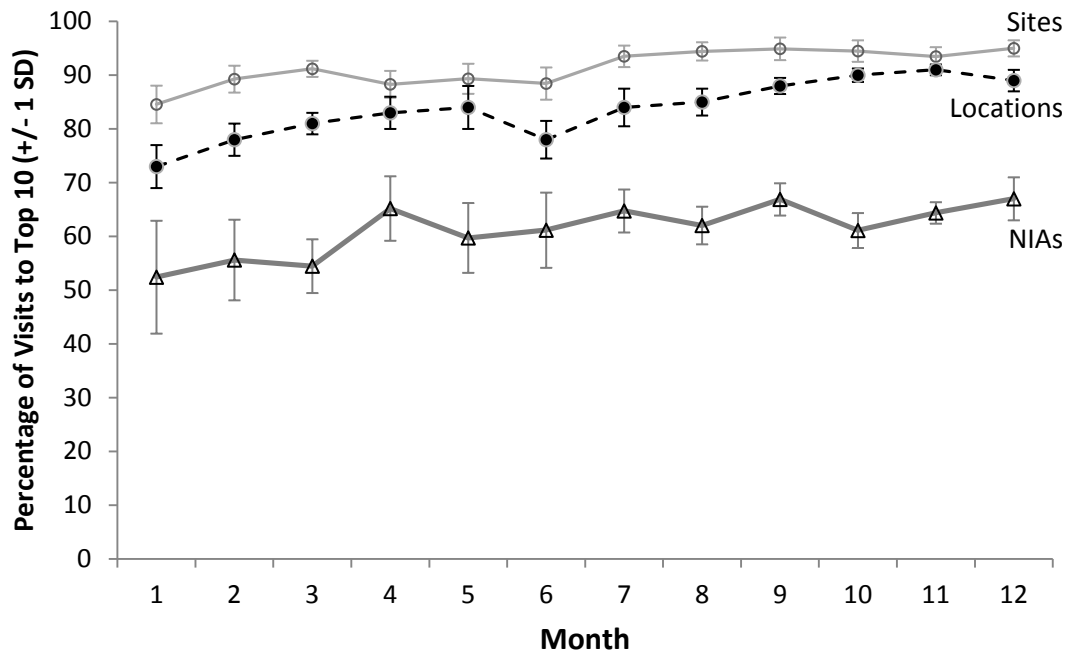
Since virtual revisitation behaviors have been examined, I now turn my attention to physical location revisiting patterns. There was high likelihood that the location where users accessed their phone was previously visited (*Location Revisitation Rate* = 90.6%). The 24 participants in the study revisited a small set of physical locations at about the same rate as virtual locations (e.g., sites, reported above at 90.3%) though the variance between users was higher for physical location revisiting ( $SD = 8.63\%$ ,  $Min = 61.3\%$ ,  $Max = 96.2\%$ ). Of course, the level of granularity of the location measurement is coarse; users could be accessing their phones in a number of settings within a given radius and still record under the same Cell ID. Still, the localities captured here reveal physical traversing following patterns similar to web localities. Even the distribution of visits and revisits to places is similar to those for sites and NIAs (Figure 17). Most locations are visited once; however, a majority of smartphone interactions occur within a small subset of places. The high revisitation rate reflects a large number of visits to users' top three to five locations similar to NIA and site revisit rates. Location revisitation rates were higher during the academic year compared to summer months, most likely because 91% of our users lived on campus during the former time period. Though smartphones are likely used in more diverse contexts compared to PCs, we found that users frequently revisit the



**Figure 17. Percentages of resources and locations by their number of visits.**

same places to interact with their phones. Across both physical and internet environments and platforms in the latter, users became more concentrated with time.

Users' unique sets of top ten sites, NIAs, and locations most frequently visited within each month of the study were analyzed as a function of experience. Figure 18 shows the trajectories for these most frequently visited resources. Most unique resources and locations were toured during the first month of the study immediately after receiving their phones. At the beginning of the study, users accessed a wider range of sites, but varied highly from each other ( $M = 85\%$ ,  $SD = 7\%$ ). Conversely, the top 10 web sites made up a larger percentage of total usage with the smallest variance at the end of the



**Figure 18. Top 10 sites, NIAs and locations for each user visited within each month.**

study compared to other months ( $M = 97\%$ ,  $SD = 3\%$ ). Similar patterns were seen for NIAs and locations. Though the type of sites, NIAs and physical locations visited differed across subjects, their concentration on a niche of favorite resources and locations to use their devices increased with time. User differences in this pattern diminished as a function of experience.

#### **5.4. Models of Physical and Virtual Visitation**

Though large diversity was found between smartphone users in the above analysis, several aspects of smartphone usage were remarkably similar across users. Namely, most users settled on a core niche of resources that were revisited very

frequently. Additionally, users accessed their iPhones in a small cluster of locations. In this section, I present these similarities by developing simple models.

Many sources of information have been characterized as following the power law (Anderson & Schooler, 1991). For instance, the recency, frequency, and spacing of words in New York Times headlines, spoken words to infants, and electronic mail predict the information that can be found in future iterations of these communications according to a power function. Equation 7 defines the power function:

$$P = At^{-b} \quad (7)$$

where  $A$  and  $b$  are constants and  $t$  represents a period of time. Of course, Anderson and Schooler's (1991) main interest was not characterizing these environmental sources (e.g., content within the New York Times) per se, but their relationship to human memory. They found that the exponent  $b$  is the rate by which information is forgot when assessing retention or the learning rate as a function of practice. Simon (1957) used this constant to show the probability of an item being repeated in the environment was proportional to its previous frequency of occurrences. One benefit of the power function above is that it can be expressed linearly under a logarithmic transform such that:

$$\log P = \log A - b \log T \quad (8)$$

The power law, or slight variations of it, has been applied to understand how users browse the web on a PC. Pitkow (1997) used the power law to predict web page visits based on frequency and recency to enhance caching.

One such derivative, the Zipf distribution, has been used often to characterize social and natural systems. Zipf (1949, 1972) demonstrated that the frequencies of words in a book plotted against their statistical rank on a logarithmic scale yielded a straight line with a slope near -1. Zipf's law has also been applied to account for a wide range of natural phenomena (e.g., Mandelbrot, 1977) such as words in human languages, city sizes, earthquake magnitudes, and income along with internet characterizations including site visits, network activity, and caching patterns. Zipf's law can be expressed mathematically as a power law function in that the probability of attaining a certain size of  $r$  is proportional to  $r^{-\alpha}$  where  $\alpha$  is close to unity. For this study, the number of visits to NIAs, sites, and locations (denoted as  $F$ ) relative to its rank (denoted as  $r$ ) are inversely proportional such that

$$F \sim c r^{-\alpha} \quad (9)$$

where  $c$  and  $\alpha$  are constants. Taking the logarithm of each side of the equation returns

$$\log F = \log c - \alpha \log r \quad (10)$$

and the slope of the rank coefficient  $\alpha$  is near 1 if Zipf's law holds (Adamic & Huberman, 2002).

As shown in Figure 19, Zipf's law applies quite strongly to internet use on the smartphone and the locations where these devices were used. The log-log plot shows a nearly linear relationship with slopes all less than -1 with good fits to the data ( $R^2$  values over .95). These steep slopes verify what was reported above: there exists a niche of very popular resources and locations that generate the most visits and less popular resources

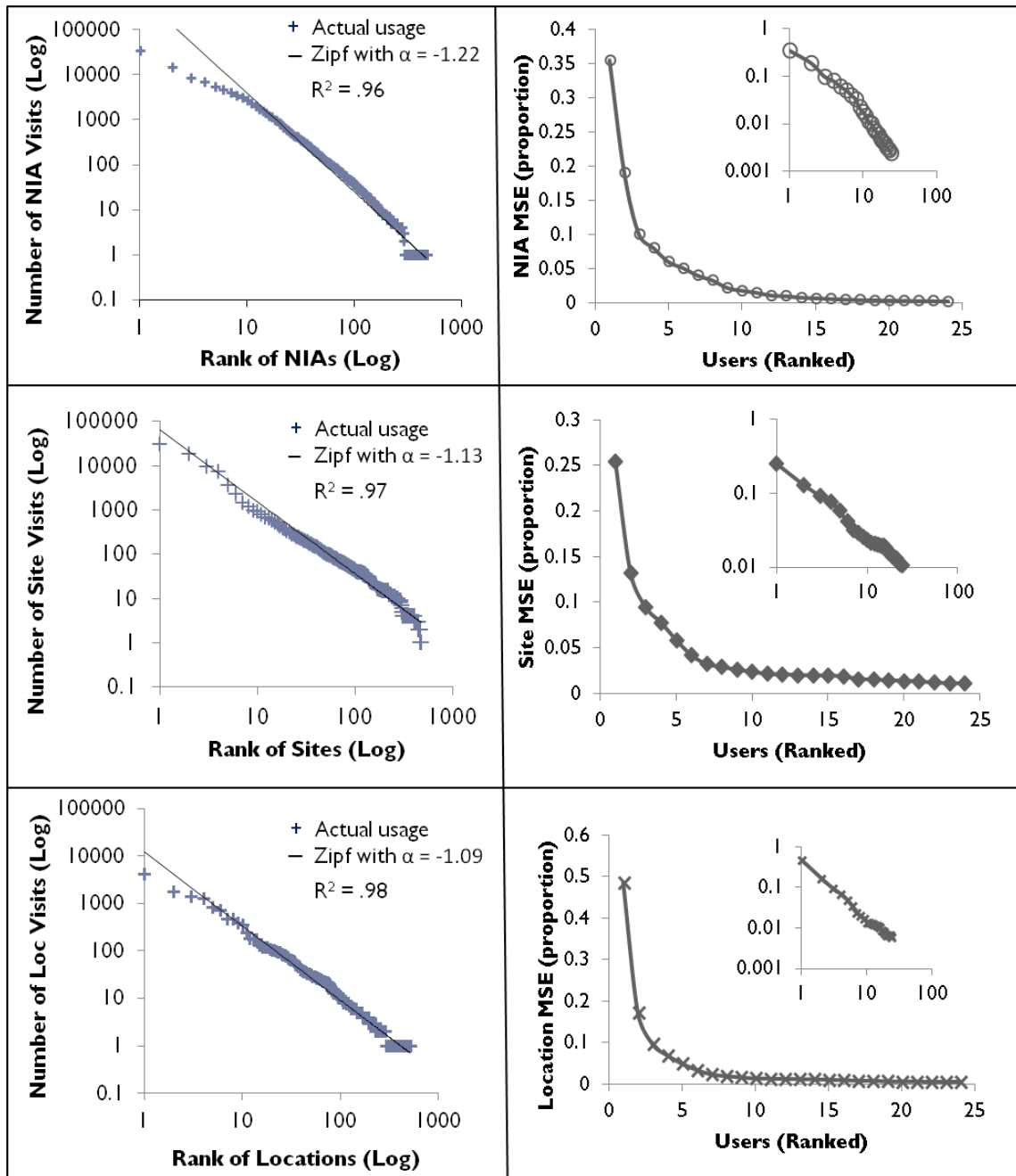


Figure 19. Fitted Zipf distributions for ranked NIA, site, and location visits (left column) alongside the corresponding mean square error (MSE) of users' actual distributions against the Zipf distributions. The insets within the latter are log-log plots.

and locations produce substantially decreasing amounts of visits. The slopes less than -1 are likely due to the low number of diverse resources and locations accessed. Also, note the convexity in the actual data distributions for both NIA and location visits. These differences between the actual data and model data suggest a more even distribution of visits to the most popular resources and locations. That is, instead of one element dominating the next most popular resources, a small set of resources are used very frequently.

How well does Zipf's law apply across all participants? Each user's data were individually fit to a Zipf model. The actual distributions of each user were compared with the model distribution using mean square error (MSE). As can be seen in the right column of Figure 19, the steep slope in visits to resources and locations as a function of popularity applied for most users. Only a few users accounted for most of the variance from the Zipf fit to NIA, site, and location visits and most users' visit patterns followed a Zipfian drop in usage based on popularity rank. Interestingly, the relationship between visits and user variance calculated by MSE is inversely proportional such that the highest ranking individuals (i.e., the top few most variant users) disproportionately explain most of the error between the model and actual data. Most users' data, however, fit Zipf's law quite well as indicated by the long tails. In other words, these slopes seem to follow a Zipf-like distribution as indicated by the right column in Figure 19 and the inset plots therein.

## 5.5. Exploring User Diversity

Until now, I have described empirical patterns of behavior associated with internet use via smartphones and the physical locations where users interacted with their smartphones. In this section, I further explore user variance reported above to characterize smartphone *users* in their visiting patterns. Many of the variables being examined have been introduced above, but are defined briefly here for the convenience of the reader:

- Total Internet visits: Total number of site and NIA visits.
- NIA-to-site index: Ratio of app sessions to browsing sessions. Values higher than 0 reflect users that use applications more than browsing to connect with the internet. Users yielding values less than 0 show the opposite pattern.
- EntroDiversity Index: Reflects the variability of application use. Higher scores represent more sporadic use across a wider range of applications. Lower scores represent narrower use of a smaller set of applications or more concentrated use of fewer applications.
- Number of URLs per session: The number of pages per web browsing session.
- Site recurrence: The site revisitation rate.
- Page recurrence: The URL revisitation rate.
- NIA recurrence: The NIA revisitation rate.
- Location recurrence: The location revisitation rate.
- Query rate: The proportion of browsing sessions that included at least one search.
- Mean number of queries per session: The mean number of queries made per web browsing session that included at least one search.
- One-URL session rate: The proportion of browsing sessions that consisted of visits to only one URL before closing.
- Duration Per Session: The mean duration per browsing session.



- Site visited once: The proportion of all sites that consisted of sites visited once.
- NiAs visited once: The proportion of NiAs visited once to all applications visited.
- SUS Score: Scores on a 0-100 scale reflecting perceived usability of their iPhones. Higher scores reflect higher perceived usability.

Since a primary interest is to understand user differences in their reliance on browsers vis-à-vis NiAs (see research question 3), I explore how users that differ in their relative use of each as measured by NIA-to-site indices. For clarity and reasons explained below, I call users lower in NIA-to-site indices (and higher total internet use) **Pioneers** because, while they visited NiAs (i.e., “native territory”), they also frequently explored the web. Those with higher indices are called **Natives** because they largely avoided exploring the web on their browsers, but accessed resources designed specifically for their smartphone technology. I keep the index as continuous intentionally to avoid strict compartmentalization of users into two user types. Instead, I imply that users at each end of the spectrum can manifest behaviors at the other extreme, though perhaps not as frequently.

Pioneers’ larger reliance on their web browsers resulted in higher site ( $r = -.54, p < .01$ ) and page ( $r = -.44, p = .02$ ) revisitation rates. There was not a significant correlation between NIA-to-site index values and NIA revisitation rates ( $r = .09, p = .69$ ). Interestingly, however, there was evidence that Pioneers revisited physical locations at a higher rate though not quite reliable at a .05 alpha level ( $r = -.26, p = .10$ ) and used their phone in fewer unique localities across the entire study period ( $r = .29, p = .08$ ).

When the browser was accessed, Natives tended to use it in conjunction with NIAs and for quick searches. As NIA-to-site indices increased, the proportion of browsing sessions that consisted of only one URL that loaded before the session ended increased as well ( $r = .51, p = .01$ ). 63% of these followed access to another NIA. For instance, many of these browser visits occurred directly following the Mail NIA suggesting that Natives were following a link in an email message. Interestingly, Natives also yielded higher query rates (proportion of browsing sessions that included at least one search) when they did access their browser ( $r = .41, p = .02$ ). The strength of this correlation increased when we removed browsing sessions that directly followed use of another NIA ( $r = .55, p < .01$ ). In other words, when Natives launched Safari from their springboards there was a high probability that a search would be issued.

Pioneers accessed their browser more from their springboard and in isolation from other NIAs. This led to different browsing patterns. As the NIA-to-site index decreased, users tended to access more pages per session ( $r = -.39, p = .03$ ) leading to each session lasting for longer periods of time ( $r = -.37, p = .04$ ). Though they yielded fewer sessions with queries (lower query rates), Pioneers averaged more queries per browsing session ( $r = -.37, p = .04$ ). Thus, when Pioneers used search they tended to search more within the same session. This resulted in more new content consumed reflected by higher unique site vocabularies per session ( $r = .69, p < .001$ ). Clearly, Pioneers relied on their web browsers for repeated visits to pages and then ventured to explore new information.

Natives used NIAs differently than Pioneers. First, they tended to spend less time on each NIA launch, though not quite significantly different ( $r = .30, p = .06$ ). They did

not differ from Pioneers in the installation of new NIAs. Thus, since Natives and Pioneers yielded similar NIA revisitation rates and vocabularies, it appears differences in use of NIAs is influenced by other factors. Indeed, the coarse location revisitation rates we report above highly correlate with NIA revisitation rates ( $r = .63, p < .001$ ) suggesting external stimuli likely prompt routine (or habitual; Oulasvirta et al., 2011) use of NIAs. Still, user differences characterized here show differences in how NIAs are accessed. Natives accessed more NIAs once without any revisits ( $r = .41, p = .02$ ). Interestingly, Pioneers uninstalled more NIAs ( $r = -.34, p = .04$ ). Finally, Natives were more consistent in their interactions with NIAs, while Pioneers were more erratic. EntroDiversity scores inversely correlated with NIA-to-site indices calculated for each user ( $r = -.49, p = .01$ ). Natives use NIAs briefly, more consistently and leave unused NIAs deserted on their springboards. Pioneers, on the contrary, were less predictable in how they used NIAs, spent more time on each NIA when launched and uninstalled more NIAs. Interestingly, Pioneers also perceived their iPhones as more usable ( $r = -.44, p = .02$ ).

## Chapter 6

### DISCUSSION

The goal of this study was to characterize internet use on smartphones using a deliberately naturalistic and longitudinal methodology. Indeed, because our logger unobtrusively collected interaction data from users' smartphones in ecologically-valid environments over a substantial period of time, I submit the behaviors examined in this study were particularly realistic. All of the users studied were not previous smartphone owners. Therefore, the findings above represent statistics and patterns taken from novice users as they gain experience with their device. This control also revealed similarities in some trajectories of usage and how users optimized their device usage as a function of experience. At the broadest level, I found differences between [1] technologies (the PC and the smartphone), [2] platforms (NIA and browser) and [3] users (Pioneers and Natives) in how the internet was accessed over time. I discuss these findings in turn before design recommendations are provided.

#### 6.1. Characterizing Web Use on Smartphones

Clearly, the web browser is not as fundamental to smartphone use as it has been reported to be with PCs. The *highest* frequency user in the current study averaged fewer URL visits per day (20) compared to the *lowest* frequency user in the most recent PC study published in 2008 (24.9). This difference between technologies was both because the browser was accessed less often and, when it was accessed, it was for shorter periods of time. Indeed, smartphone browsing sessions were roughly three times shorter than PC

browsing sessions in both duration and pages visited. Similar to the major usability problem reported for the PC last decade, mobile browsing suffers from long page loading delays and this was noticed by our users. Additionally, NIAs accounted for a large number of visits to the internet.

The PC web browser has been characterized as a recurrent system due to URL revisitation rates higher than 50% (e.g., Tauscher & Greenberg, 1997). Because of this classification, supporting users' revisits to previously accessed content is essential as evidenced by the long thread of HCI research on this topic. The page revisitation rate computed here for the smartphone (25%) is much lower than any page revisitation rate computed previously for the PC. Indeed, over 75% of all pages were visited only once or twice. The pages that were revisited more than a handful of times were primarily top-level pages used as a portal to get to new content. Other revisits occurred because of how the browser functions when launched. The last page open before the browser is closed is loaded and displayed when opened again. Thus, many revisits were continued browsing sessions after an interruption (recall Table 5). Longer-term revisits generally occurred through search which is an indicator of the need for enhanced interface support to reaccess desired content on browsers (Tauscher & Greenberg, 1997). The surprisingly low use of bookmarks overall also provides evidence supporting this conclusion. Since we examined web logs, it is unclear if the low page revisitation rate is due to browser limitations instead of their actual information needs and desires, though our user reports suggest the latter. This chicken-and-egg problem should be addressed in future research using other observational techniques (Byrne et al., 1999; Cockburn & McKenzie, 2002).

Though URLs were not revisited often on browsers, sites (i.e., domains) were revisited at a higher rate than reported in any of the previous PC studies. Only a handful of sites accounted for most of the page visits for each user. At the beginning of the study, users toured a large number of unique sites and, subsequently, developed large vocabularies. These vocabularies shrunk considerably as most of the sites and pages visited were not revisited. Instead, users settled on a niche of sites that were revisited frequently throughout the study, though unique pages within these sites were commonly accessed. Clearly, a small set of domains provide users access to a much larger number of diverse pages suggesting designers should provide easier access to these portals (i.e., avoid making users search and navigate to these home pages). For example, a future browser design could add a bookmarks bar within the browser bar similar to recent versions of Safari for larger computers. Additional suggestions are provided below.

## **6.2. NIAs and Browsers: NIAs Dominate, But the Web is Not Dead**

The distinction between computing technologies in volume of use became more pronounced with experience as browser use gave way to higher NIA activity. NIAs consumed most visits to the internet, especially after users became more experienced with their devices. These visits were short and concentrated to a relatively stable vocabulary of NIAs that were frequently revisited. When our users first received their smartphones, they seemed to apply mental models developed from computer use and relied heavily on their smartphone browser for touring a large number of sites and accessing these sites often. However, instead of bookmarking these sites within their browsers, users installed NIAs that “stuck” to their vocabularies for longer periods of

time. Eventually users access more unique NIAs within a given month compared to sites within the browser. In a sense, smartphones such as the iPhone are designed around “NIA bookmarking” as a primary means to access resources. Perhaps one reason users did not bookmark within the browser was because they installed the NIAs they wanted to revisit. NIAs “bookmarked” by users through the App Store are added to springboard pages and can be personalized by users without the structure of the browser interface. Indeed, users install and personalize NIAs on their springboards based on their perceived future need for mobile services similar to how URLs are bookmarked on a PC browser. It appears, based on this research, that the former can be improved. Springboard real estate was important and users frequently “guessed wrong” when predicting future information needs judged by the number of uninstalls. Or, more likely, they “guess right” in predicting their future information needs and install NIAs that they think will be useful; however, they also quickly discover the NIAs installed are not useful or entertaining. Many NIAs were not given a large number of opportunities to show these characteristics. Similar to content on the web, NIA visits followed a Zipf structure and a large number of NIAs that were installed by users were subsequently uninstalled or left on a springboard page after only one or two visits. Of course, this may not reflect poor predictions and could represent the fact that some NIAs have a short shelf life. Games, for example, seem to fit into this mold. Still, apart from search, NIAs with the right capabilities are more likely to be revisited than sites on the browser and users relied heavier on NIAs over sites as a function of experience. The clear message to web designers is to point users to installing the NIA to increase the probability of revisits.

Even if a NIA is installed, our study suggests that it is difficult for new resources to become part of users' niche of NIAs that are regularly revisited. Interestingly, the number of unique NIAs visited within each month did not increase much across the entire length of this study, though new NIAs were installed. This suggests total NIA vocabulary space is somewhat fixed for users; as new NIAs are added, others are either uninstalled or left on springboards with low likelihood of revisits. Indeed, EntroDiversity scores revealed this predictability in terms of information entropy and showed that users access their vocabulary of NIAs in less sundry ways as a function of time. In other words, the visits to users' stable set of NIAs were concentrated on a smaller portion of their vocabulary. Clearly, use of the internet away from the desktop still resulted in recursive behaviors that became more routine and predictable with time.

Taken together, NIAs were used more frequently to access a stable set of resources in particular contexts. Web browsers became a tool that afforded searching for and consuming new and dynamic information with low likelihood of repeated visits. Adding NIAs required users to predict their information and entertainment needs, install the NIA without assessing its usefulness, and take up springboard real estate while the web browsers were used more ad hoc. These findings seem to point toward a continued need for web browsing, even as the number and use of native applications continues to soar. Across all users, when the browser was accessed, it seemed to be used for six types of activities:

- Discovery (don't know what I need, but Google can help)
- Know what I need, but don't need it for long
- No native application available to access desired content



- Native application is available, but doesn't have the capability needed
- Supporting use of another native application (e.g., link from an email message)
- Interrupted browsing (e.g., Table 5)

The mobile web seems to have similar usability problems (e.g., page loading) reported for the PC-based web last decade. Based on the current state of the mobile web, it appears the web is not dead, but a secondary platform to access the internet that affords ad hoc traversing to find and use needed resources without the cost of installing a NIA.

Search was a hallmark of browser use and became more concentrated across all browser sessions as a function of experience. This trend is opposite of previous reports associated with the effect of experience on searching with a PC web browser (Cothey, 2002). Interestingly, internet searches were barely issued through NIAs though most users had capabilities on their devices to do so. It is unclear what drove the extremely low use of the Google Mobile NIA; however, the long page loading times recorded in Tossell et al. (2010) is a likely culprit. Though the web is not dead, the URL address bar in its current form seems to be dying. Users did not appear to directly access resources using this capability. Since the time of this study, however, a menu based on initial inputs from user text entry is displayed showing titles of sites that have been visited previously and start with the same letters. This could increase use of the URL address bar to afford revisits while the neighboring search bar could afford exploring.

In contrast, NIAs became the default mechanism to access resources on smartphones and ultimately the primary vehicle to access the internet. Users mentioned that if an NIA was available for their need, they would prefer accessing resources through

it instead of the browser. Many times these visits were prompted by contextual cues. For instance, one user admitted to opening the same two or three NIAs before class even though he knew there was nothing new to see. Instead of passing the time with browsing new information, users seemed to engage in predictable routines evidenced by low EntroDiversity scores and the proportion of used NIAs to total NIAs presented above and additional research (e.g., Oulasvirta et al., 2011).

Even visiting patterns of physical locations where smartphones were used followed highly recursive patterns similar to site and NIA revisits. Users regularly used their phones in a small subset of locations resulting in a high revisitation rate (90%) similar to revisits to locations on the internet. Of course, this could be due to the unique population of users being studied (79% of the students that participated in this study lived on campus during at least a portion of the study period) and other populations should be assessed in future research. Still, this high degree of similarity between physical and virtual revisitation was somewhat surprising, though the three primary revisitation rates were each measured at a coarser level of granularity. For instance, the actual pages visited within NIAs (e.g., News Feed, Profile and Friends List all within the Facebook application) were not captured by our logger. Similarly, the settings (e.g., a classroom, dorm room, etc.) where iPhones were used were not recorded; only the Cell IDs which could provide service to users in a number of settings. In contrast, we did capture browser navigating at a lower level of granularity; users explored new content within regularly accessed sites. Like many other social and physical phenomena, interactions with smartphones in the real-world follow a Zipf structure. This suggests, even for ubiquitous computing, interface support for reaccessing each user's unique set of frequently visited

internet resources is essential. Additionally, since users revisit the same physical locations frequently, predictive systems should leverage this information (i.e., location) to preload frequently visited internet content (e.g., update Facebook, download emails, load a top-level page on to browser, etc.) to attenuate problematic page loading delays. For instance, since users settled on a handful of NIAs and sites, future capabilities should allow users to set their devices to automatically load these resources prompted by their cell ID location or when content within those resources change. Future mobile web systems, leveraging HTML5, should allow users to be alerted when content within pages change (e.g., a new item is available in Craigslist, a blog has been updated, etc.) to reduce unnecessary web traversing. Even without an alert, springboards should provide an indicator showing which highly revisited web resources have new information similar to functionality provided by some NIAs.

Fully automatic systems capable of exploiting each user's history of usage and context would not require users to set a customization for this capability to be enabled (e.g., similar to the Windows Start Menu; see Chen & Kotz, 2000 for a review of context-aware systems). Correlations between revisit rates to these coarse virtual (internet) and physical locations shows users behave similarly across environments. User traversing patterns are stable across environments suggesting patterns logged in the real-world can inform internet use patterns and vice versa.

Although longer-term research leveraging context to predict resource use could be beneficial to attenuate HCI issues such as page loading delays, it appears that simply exploiting user history can be just as beneficial. This is especially the case for more

experienced users. Even simple predictive models, such as frequency, can attenuate some of the problematic page loading delays. Simply preloading and updating the 3-5 most frequently used NIAs and sites periodically could be a beneficial first step. Because users visit the same localities to access their phones (and then go on the internet), it appears users develop fixed routines with their smartphones prompted by where they go in the real-world. In other words, this temporal precedence seems to suggest contextual settings prompt habitual NIA use as suggested in previous research (Oulasvirta et al., 2011). Of course, the large variance between users suggests that implementing predictive models and context-aware systems should be done cautiously and allow for user-driven personalization based on each person's unique needs, desires and other contextual influences.

### **6.3. User Differences in Mobile Internet Use**

Previous studies have found differences between smartphone users of several orders of magnitude (Falaki et al., 2010). The present thesis corroborates this finding and suggests there is no average smartphone user. However, the large differences between users can be characterized building on a growing line of empirical research in HCI. Catledge and Pitkow (1995) described differences between PC web users based on browsing patterns. One type, "Serendipitous Browsers" did not yield repeated sequences of URL visits. "Searchers," in contrast, were repetitive in short navigation sequences. Teevan and colleagues (2004) labeled users at these extremes as "Filers" and "Pilers." Filers designated explicit locations to organize their electronic information. Pilers, on the other hand, were unstructured and used different search strategies to retrieve their

electronic information. Extending findings in education, Ford and colleagues (2002) found that “Holists” issued more exploratory searches and valued serendipitous encounters with new information. In contrast, “Serialists” follow a linear pattern to learning and navigate on the web in a more sequential manner. White and Drucker (2007) similarly described PC users as “Explorers” and “Navigators.” The former prefers undirected browsing and discovery of new information. The latter prefers rapid access to target information and performed more directive searching to desired content.

From the results above, characterizations were developed of mobile users at two ends of a similar behavioral continuum. At one end, Pioneers relied on their browsers more and these interactions yielded visits to more diverse content, longer sequences of URLs per session, more searches within each browsing session, and higher rates of page and site revisiting. This larger consumption seemed to reflect that these users “settle” on a small subset of favorite resources on the browser very frequently. However, they also continue to pioneer in these browsing sessions with search and visiting of new content for more sundry reasons. This behavioral pattern manifested somewhat in NIAs as well; perhaps exploring more within NIAs reflected by longer visits and clearing out room on their springboards for new NIAs. Pioneers also revisited physical locations at a higher rate and used their iPhones in fewer unique localities. Clearly, these users are most similar to “Serendipitous Browsers,” “Holists,” “Explorers” and “Pilers” in that they consume more diverse amounts of information, interact for longer sequences when on the web and traverse in more localities in the real world.

On the other end, Natives did not access the internet as much and did not venture outside of applications native to their device to content on their web browser. They are most similar to “Searchers,” “Filers,” “Serialists,” and “Navigators” in this regard and because they rely on shorter navigation sequences using search when on their browsers. Web browsers were sparingly used to discover new information; though they were used directly after other NIAs. Natives yielded web sessions shorter in terms of both pages visited and duration. Similarly, NIA launches were for shorter durations as well.

Even with a small sample size, the within-user variances measured by ED, differed across users. Some users accessed their NIAs erratically relative to others that operated their devices more predictably. The former were aligned with Pioneers corroborating previous research suggesting that Explorers browse in more sundry patterns (White & Drucker, 2007) and Filers lack of strict organization strategies associated with their electronic content. The latter operated their phones more predictably using a concentrated number of NIAs. Natives were more focused and directed in their interactions similar to Navigators’ linear searches to targeted information and Filers’ structured organizational strategy.

How can these user differences be explained? First, the distinctions could reflect patterns explained by differing cognitive styles identified in previous learning and HCI studies (see Witkin, Moore, Goodenough & Cox, 1977). Because stable behavioral differences, similar to differences between *holists* and *serialists*, manifested across virtual and physical traversing in our study, our smartphone users differed because of how they varied along this spectrum. Second, these differences between users could have

manifested because of the types of contexts visited and user goals within those settings (see Lee et al., 2005). Natives seemed to actively use their smartphones more as tools to accomplish short information needs. Pioneers seemed to use their smartphones actively and passively for both utilitarian and hedonic reasons. Pioneers may have developed more habitual routines as they revisit more of the same locations to access their phones. Since they do not visit more physical locations, perhaps they compensate by traversing more on the web to content previously settled on along with additional explorations. Third, these differences between users could be driven simply by volume of overall usage and perceived usability. Because Pioneers accessed their devices more and perceived it as more usable, their amount of information needs could have triggered more diverse behaviors in addition to routine use. Finally, our results could be driven by interactions between these three factors. These hypotheses should be tested in additional research.

#### 6.4. Design Implications

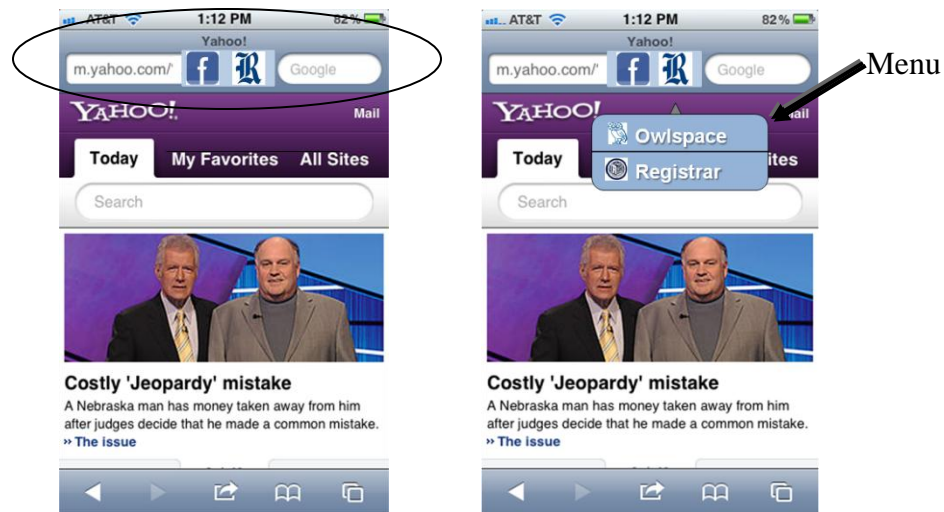
Findings from this dissertation suggest several ways to better support smartphone users. Even though these recommendations are directed toward users at the ends of the continuum, I submit that each suggestion would be beneficial for all users. In other words, users along the entire spectrum can display both types of behavioral patterns and gain from these enhanced designs. **Pioneers** could benefit from capabilities previously reported for similar types of users along with several suggestions for mobile space:

- **Optimize content for mobile browsing:** Content designers should not ignore designing a usable mobile site for web browsers. Many new smartphone users rely

heavily on their browsers to access the internet and find information to be bookmarked via installing the NIA. Additionally, the mobile browser continuously provides a vehicle for many users at all experience levels to access new content with low likelihood of revisiting. Optimized content for mobile browsing should assume users are new and provide clear “knowledge-in-the-world” to support first-time interactions along with clear links to add bookmark through installing the NIA, adding the site tag to a springboard, or the top browser bar.

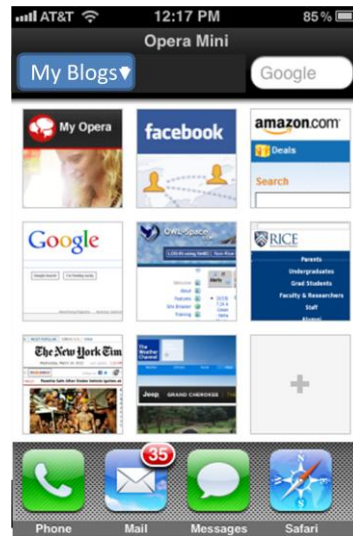
- **Better design of mobile browsers:** Results of this research suggests that the real estate used for the URL address bar can be better exploited for personalized access to favorite content. Since many of the sites and pages revisited were top-level pages (e.g., home pages), designers should provide easier access to these portals to avoid unneeded typing, searching, and page loading. For instance, the Google search bar could double as the URL address bar for pasting URLs and for autocomplete functionality to revisit pages. The free space on the top browser bar (no longer occupied by the URL address bar) could be used to access top-level sites based on each user’s most visited sites or adaptable for users to quickly add and edit based on their perceived needs. Less drastically, the length of the URL address bar could be decreased to make room for users’ top two sites (Figure 20). For sites with a short tail (such as the institutional sites in Figure 7), these personalized buttons could display a menu with the most visited sub-level sites when selected.





**Figure 20.** An example of an enhanced browser bar that facilitates easier access to top sites. Menus can be displayed for sites with high page revisits (e.g., institutional sites) to the sub-level pages that are frequently accessed.

- **Personalized browser pages on springboards:** Beyond the browser bar, springboards can also be designed to better facilitate reaccess to personalized niches on the mobile web. Instead of every springboard looking the same, one springboard can provide users with indicators of what has changed within their favorite sites and links to that content (transporting directly to this content over top-level pages and/or search results). The Opera mini browser home page provides an example of such a display for one springboard page (Figure 21). Additional space on this web springboard could be leveraged for personalized suggestions of other web resources similar to Amazon for books and Netflix for videos.



**Figure 21. An example of a personalizable “Web Springboard.” The Opera mini browser provides users easy access to top sites from its home page. A similar approach could be used for a springboard designated for browser use.**

- **More intelligent springboards:** It seemed very evident from this research that springboards can be designed more effectively to support discovering new NIAs as well. This could take several forms. For example, user profiles could be developed from NIA and browsing history to provide suggestions based on content visited. Smartphone designers could give users the option to designate one springboard page for recommender capabilities (e.g., Figure 22). This springboard may allow users to try NIAs before they permanently install or buy applications. In other words, a recommender system would install suggested NIAs for trying out instead of relying on the user to browse, install, and try out the app for him- or herself. Most users had room on their springboards for such a capability. We recommend, especially for Pioneers, a springboard page designated for “Try-before-you-buy” NIAs already installed based on previous usage. Part of this springboard space

could be used for browser access. For instance, an auto-bookmark mechanism to provide links to a top-level site from a user's unique history could be leveraged similar to the Windows Start Menu. Perhaps even a static top-level page with links to sub-level resources would be beneficial for more efficient interactions. Another suggestion is to support users when they visit new locations. Since location visits followed a Zipf distribution, smartphone users accessed their phones in a large number of locations very infrequently. Springboards such as Figure 22 could better support users when they visit a location in their long tail. Using information about what NIAs others have used at particular Cell IDs, users could have quick access to NIAs potentially important for that location and allow users to explore and find new capabilities of interest.



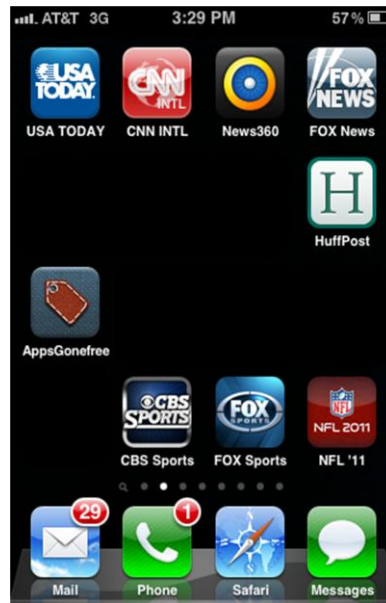
**Figure 22.** Examples of intelligent springboards that leverage users' history of application installs (left) and location information (right).

**Natives** could also benefit from design features to help them get to desired information more quickly and better structure their springboards.

- **Predictive systems:** More predictive capabilities are needed to attenuate mobile HCI problems mentioned above such as page loading and awkward text entry. Activity gathered from other devices can be useful to this end (Kane et al., 2011). According to this study, each user's history of web use on their smartphones could provide ways to predict likely site destinations. Most overall internet use was dominated by a relatively small set of sites and NIAs. Using contextual information and most frequented sites to preload pages (e.g., top-level pages) would be beneficial to avoid long page loading delays.
- **Web search from the springboard:** When Natives accessed their browsers, it was usually for quick searches. Queries were conducted most often on the web. It may be beneficial to offer a vehicle to search from the springboard. For instance, instead of capabilities to "search your iPhone" from the springboard, it may be more useful to have a "search the web" bar right on the springboard main pages. Of course, it appears giving users options to do both would be optimal. It is also unclear why the "I feel lucky" option (i.e., giving users the option of being transported directly to the URL of the top search result without viewing a search results page) is not available for mobile browsing. Allowing users to skip unnecessary page visits and thus page loading delays seems more important for the current mobile web than for the stationary web given the long page loading delays (see Tossell et al., 2010). I submit that giving users this option, especially those that use their devices more like

Natives, would be beneficial. Since most web searches took place from the browser bar, this location is most important for the proposed functionality. However, these options would also be beneficial for quicker access from the springboard and search bars within web pages.

- **Identify Springboards:** Across the entire study, NIA vocabularies were generally higher than what could fit on one springboard page. Natives did not uninstall NIAs much compared to Pioneers. Also, they used more NIAs just once. Taken together, it seems like springboards could be designed to more effectively support revisits. Perhaps one way to do this is giving users options to identify each springboard. This could be done in a number of ways. For instance, future smartphones could allow users to tailor the wallpaper on each springboard page (e.g., an airplane behind a springboard page of travel applications or a purse behind shopping applications. Another option is for smartphone springboards to follow what many PCs and laptops allow on their Desktop screens and provide an option for users to turn off auto-positioning of native applications. This flexibility could afford more distinct looks between springboards for easier identification, ease searching for a native application on a particular springboard, and perhaps support more personalization of native applications into meaningful clusters. Figure 23 provides an example of this suggestion.



**Figure 23. An example of more personalizable springboards for enhanced identification of NIAs.**

Similar to work before this one (e.g., White & Drucker, 2007), I have intentionally focused on users at each extreme in these recommendations. Of course, most users are likely not at the ends of the continuum, but somewhere in between. These implications, thus, apply broadly and likely would benefit all users.

## **6.5. Limitations**

Of course, these findings should be generalized with caution because of several limitations in the current study. Foremost, our small sample size does not represent the entire range of the millions of smartphone users around the world. The students at Rice University who made up our sample are also different from the general population in other ways as well (e.g., age, education, etc.). However, the number of users analyzed

here is roughly equivalent and perhaps more diverse than many of the previous studies that have informed our research. Second, I only examined the use of iPhones. Future research should assess how the results presented above generalize to other devices. Third, the PC studies are slightly older and we suspect current PC web use has changed (at least for smartphone owners). It is unclear how use of a smartphone with a PC impacts the latter. Many users also have other types of devices not assessed in this dissertation. For instance, laptop computers and tablet computers provide capabilities that are less stationary than the PC on devices that are larger than smartphones. Future studies should assess how users access the internet from a range of technologies to get a more complete picture.

## **6.6. Conclusion**

Bearing these limitations in mind, this study contributes empirical characterizations of internet use on smartphones. Behavioral patterns associated with browsing, NIA use and physical location visits were examined and user differences were explored. As the world move towards Weiser's vision of ubiquitous computing, designers of continuously-available internet technologies must account for what users do on the internet. Most research in mobile space suggests that smartphones are used in sundry environments and that usage patterns are extremely dynamic and unpredictable. This dissertation shows that humans are the proverbial "creatures of habit" visiting the same general locations both in the real world and on the internet at a high rate. Seemingly complex and noisy phenomena, such as the locations where users access their smartphone

and what they do on these devices and user variance, tend to follow Zipf's law similar to other social and physical systems.

Moreover, experience seemed to decrease the variance between users and led to more systematic and concentrated visit patterns across all users. New smartphone owners are somewhat erratic in their NIA and web use and tour a large number of internet resources. Over time, users become less eclectic and more systematic in where and how they use their smartphones echoing other research within learning paradigms (e.g., Tossell, Schvaneveldt & Branaghan, 2010). On the PC, experience led to more eclectic and passive usage. On the smartphone, however, variance between users in many aspects of usage decreased and usage became more active as a function of experience (e.g., using the browser to search).

Indeed, computer use away from the desktop did not lead to more eclectic and erratic usage in all respects. Instead, a small set of locations where users accessed their phones seemed to prompt repeated visits to a concentrated set of most used resources. Herbert Simon (1971) stated it best, "What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it" (pp. 40-41). This statement seems even more apropos within the ubiquitous computing paradigm and, in particular, smartphone use. As more information becomes available in more environments, humans are the agents that restrict what is being consumed. Future HCI research in ubiquitous computing, then, should not just be concerned with providing more information in more environments, but



work on providing the right capabilities in the right environments for the right individuals.

## REFERENCES

- Abowd, G.D. & Mynatt, E.D. (2000) Charting past, present, and future research in ubiquitous computing, *ACM Transactions on Computer-Human Interaction*, 7, 29 – 58.
- Abowd, G.D., Mynatt, E.D. & Rodden, T. (2002). The human experience. *IEEE Pervasive Computing*, 1, 48-57.
- Adamic, L.A. & Huberman, B.A. (2002). Zipf's law and the internet. *Glottometrics*, 3, 143-150.
- Adler, P.A., Adler, P. (1988). Observational techniques. In N.K. Denzin & Y.S. Lincoln (Eds.). *Collecting and Interpreting Qualitative Materials*. Thousand Oaks, CA: Sage.
- Anderson, C. & Wolff, M. (2010, September) The web is dead. Long live the internet. *Wired*. Retrieved November 2011 from:  
[http://www.wired.com/magazine/2010/08/ff\\_webrip/all/1](http://www.wired.com/magazine/2010/08/ff_webrip/all/1).
- Anderson, J.R. & Schooler, L.J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Bales, E., Sohn, T. & Setlur, V. (2011). Planning, apps, and the high-end smartphone: exploring the landscape of modern cross-device reaccess. In K. Lyons et al. (Eds.): *Pervasive 2011, LNCS 6696*, pp. 1-18. Berlin: Springer.

- Bangor, A., Kortum, P.T. & Miller, J.T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24, 574-594.
- Brooke, J. (1996). SUS: a 'quick and dirty' usability scale. In P.W.Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.) *Usability Evaluation in Industry*, pp. 189-194. London: Taylor and Francis.
- Buckland, M.K. (1992). Emanuel Goldberg, Electronic Document Retrieval, and Vannevar Bush's Memex. *Journal of the American Society for Information Science*, 43, 284-294.
- Bush, V. (1945). As we may think. *Atlantic Monthly*, 176, 101-108.
- Byrne, M.D., John, B.E., Wehrle, N.S. & Crow, D.C. (1999). *The tangled web we wove: A taskonomy of WWW use*. In *Human Factors in Computing Systems: Proceedings of CHI '99*, 544-551. Reading, MA: Addison-Wesley.
- Card, S.K., Pirolli, P., Wege, V.D., Morrison, J.B., Reeder, R.W., Schraedley, P.K. & Boshart, J. (2001). Information scent as a driver of web behavior graphs: Results of a protocol analysis method for web usability. In *Human Factors in Computing Systems, Proceedings of CHI 2001*, 498-505. New York: ACM Press.
- Carter, S. & Mankoff, J. (2005). When participants do the capturing: The role of media in diary studies. *Human-computer interaction institute*, Paper 123. Retrieved April 2010 from: <http://repository.cmu.edu/hcii/123>.

- Catledge, L.D. & Pitkow, J.E. (1995). Characterizing browsing strategies in the world wide web. *Computer Networks and ISDN Systems*, 27, 1065-1073.
- Chen, G. & Kotz, D. (2000). A Survey of Context-Aware Mobile Computing Research. Dartmouth Computer Science Technical Report TR2000-381, pp. 1-16.
- Cockburn, A. & McKenzie, B. (2000). What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies*, 54, 903-922.
- ComScore (2011, February). The 2010 Mobile Year in Review. *Press Release*. Retrieved January 2012 from:  
[http://www.comscore.com/Press\\_Events/Press\\_Releases/2011/2/comScore\\_Releases\\_Inaugural\\_Report\\_The\\_2010\\_Mobile\\_Year\\_in\\_Review](http://www.comscore.com/Press_Events/Press_Releases/2011/2/comScore_Releases_Inaugural_Report_The_2010_Mobile_Year_in_Review)
- Corlett, D. & Sharples, M. (2005). Tablet technology for informal collaboration in higher education. In J. Attewell & C. Savill-Smith (Eds.). *Mobile Learning Anytime Everywhere: Papers from Mlearn 2004*, 59-61. London: LSDA.
- Cothey, V. (2002). A longitudinal study of world wide web users' information-searching behavior. *Journal of the American Society for Information Science and Technology*, 53, 67-78.
- Cui, Y. & Roto, V. (2008). How people use the web on mobile devices. In *Proceedings of the World Wide Web Conference WWW 2008*, Banff, Alberta, Canada: ACM Press.
- Dey, A. & Abowd, G. (1999). *Toward a better understanding of context and context-awareness*. Atlanta, GA. Georgia Institute of Technology.

- Duchnicky, R.L. & Kolars, P.A. (1983). Readability of text scrolled on visual display terminals as a function of window size. *Human Factors*, 6, 683-692.
- Eagle, N. and Pentland, A. (2006) Reality Mining: Sensing Complex Social Systems, *Personal and Ubiquitous Computing*, 10, 255-268.
- Eagle, N. (2005). *Machine Perception and Learning of Complex Social Systems*. Doctoral dissertation, Massachusetts Institute of Technology.
- Ericsson Consumer Labs. (2010). From apps to everyday situations. Retrieved July 2011 from: <http://www.ericsson.com/fromappstoeverydaysituations.html>.
- Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R. & Estrin, D. (2010). *Diversity in smartphone usage*. In *MobiSys '10: proceedings of the 8<sup>th</sup> annual symposium on mobile systems, applications, and services*. 179-194. San Francisco, CA: ACM Press.
- Figge, S. (2004). Situation-dependent services? A challenge for mobile network operators. *Journal of Business Research*, 46, 149-158.
- Fjord. (2011). Understanding the smartphone usage of lower income customers. <http://www.fjordnet.com/fjord-insights/understanding-smartphone-usage-lower-income-customers>.
- Foltz, P.W. & Martin, M.J. (2008). Automated communication analysis of teams. In Salas, E., Goodwin, G.F. & Burke, S. (Eds.). *Team effectiveness in complex organizations*. New York: Taylor Francis.

Gallette, D.F., Henry, R.M., McCoy, S., Polak, P. (2003). When the wait isn't so bad:

The interacting effects of website delay, familiarity, and breadth. *Information Systems Research*, 17(1), 20-37.

Gilbert, A.L., Sangwan, S. & Han Mei lan, H. (2005). Beyond usability: The oobe

dynamics of mobile data services markets. *Personal and Ubiquitous Computing*, 9(4), 198-208.

Glassman, S. (1994). A caching relay for the world wide web. *Computer Networks and*

*ISDN Systems*, 27, 165-173.

Greenberg, S. (1993). *The computer user as toolsmith: The use, reuse, and organization*

*of computer-based tools*. Cambridge, MA: Cambridge University Press.

Haynes, S.N. & Horn, W.F. (1982). Reactivity in behavioral observation: A review.

*Behavioral Assessment*, 4, 369-385.

Hagen, P., Robertson, T., Kan, M., Sadler, K. (2005) *Emerging Research Methods for*

*Understanding Mobile Technology Use*, Proceedings of the OZCHI, Canberra,

Australia: ACM Press.

Hakkila, J. & Chatfield, C. (2005). 'It's like if you opened someone else's letter' – User

Perceived Privacy and Social Practices with SMS Communication. In *Proceedings of the Mobile HCI Conference*. Salzburg, Austria: ACM Press.

Haynes, S.N. & Horn, W.F. (1982). Reactivity in behavioral observation: A review.

*Behavioral Assessment*, 4, 369-385.

- Hazlett, T. (2001). The wireless craze, the unlimited bandwidth myth, the spectrum auction faux pas, and the punchline to Ronald Coase's "Big Joke": An essay on airwave allocation policy. *Harvard Journal of Law and Technology*, 2, 335-567.
- Hepburn, A. (2011). Mobile statistics, stats & facts 2011. Retrieved January 2012 from: <http://www.digitalbuzzblog.com/2011-mobile-statistics-stats-facts-marketing-infographic/>.
- Hiltunen, M., Laukka, M. & Luomala, J. (2002). *Mobile user experience*. Helsinki, Finland: IT Press.
- Holscher, C. & Strube, G. (2000). Web search behavior of internet experts and newbies. *Computer Networks*, 33, 337-346.
- Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44(3), 161-174.
- Hulkko, S., Mattelmäki, T., Virtanen, K. & Keinonen, T. (2004). Mobile Probes. *Proceedings of NORDICHI 2004*. Tampere, Finland: ACM Press.
- International Telecommunication Union (2012). Retrieved from: <http://www.itu.int/ITU-D/ict/material/FactsFigures2011.pdf>
- Jacko, J. & Sears, A. (2003). *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Mahway, NJ: Erlbaum.

- Jansen, B.J., Spink, A. & Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the Web. *Information Processing & Management*, 36, 207-227.
- Johnson, S.M. & Bolstad, O.D. (1973). Methodological issues in naturalistic observation: Some problems and solutions for field research. In Hamerlynck, L.A., Handy, L.C. & Mash, E.J. (Eds.). *Behavior change: Methodology, concepts, and practice*. Champaign, Ill: Research.
- Johnson, S.M. & Bolstad, O.D. (1975). Reactivity to home observation: A comparison of audio recorded behavior with observers present or absent. *Applied Behavioral Analysis*, 8, 181-185.
- Jones, R., Rey, B., Madani, O. & Greiner, W. (2006). Generating query substitutions. In *Proceedings of the World Wide Web Conference WWW 2006*. 387-396. Banff, Alberta, Canada: ACM Press.
- Jönsson, B., Svensk, A., Cuartielles, D., Malmborg, L., & Schlaucher, P. (2002). Mobility and learning environments - engaging people in design of their everyday environments. Retrieved November, 2004 from <http://www.certec.lth.se/doc/mobility1/>.
- Kamvar, M., Kellar, M., Pater, R. & Xu, Y. (2009). Computers and iPhones and mobile phones, oh my! A logs-based comparison of search users on different devices. In *Proceedings of the 2009 WWW: International World Wide Web Conference*. Banff, Alberta, Canada: ACM Press.



- Kane, S., Karlson, A.K., Meyers, B.R., Johns, P., Jacobs, A. & Smith, G. (2011). Exploring cross-device web use on PCs and mobile devices. In *Proceedings of Human-Computer Interaction—INTERACT 2009*, 722-735. Berlin: Springer.
- Kellar, M., Watters, C. & Shepherd, M. (2006). *A goal-based classification of web information tasks*. Proceedings of the American Society for Information Science and Technology. Austin, TX: Wiley.
- Khalil, A. & Connelly, K. (2005). *Improving cell phone awareness by using calendar information*. Proceedings of HCI-Interact (2005). Rome, Italy: ACM Press.
- Kim, H., Kim, J., Lee, Y., Chae, M. & Choi, Y. (2002). An empirical study of the use contexts and usability problems in mobile internet. *Paper presented at the 35<sup>th</sup> Annual Hawaii International Conference on Systems Sciences*. Big Island, Hawaii.
- Kivi, A.. (2007). Measuring mobile user behavior and service usage: Methods, measurement points, and future outlook. In *6<sup>th</sup> Global Mobility Roundtable*. Los Angeles, CA.
- Kjeldskov, J. & Graham, C. (2003) *A review of mobile HCI research methods*. Proceedings of Mobile HCI, Udine, Italy: ACM Press.
- Kjeldskov, J. & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *Int'l Journal of Human-Computer Studies*, 60, 599-620.
- Kukulska-Hulme, A. (2007). Mobile usability in educational contexts: What have we learnt? *International Review of Research in Open and Distance Learning*, 8(2), 1-16.

- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lee, A. (1992). *Investigations into history tools for user support*. Doctoral Dissertation, Department of Computer Science. University of Toronto, Ontario, Canada.
- Lee, A. (2010). Exploiting context for mobile user experience. In A. Dix, T. Hussein, S. Lukosch, & J. Ziegler (Eds.), *Proceedings of the First Workshop on Semantic Models for Adaptive Interactive Systems (SEMAIS)*. Virginia: ACM Press.
- Lee, I., Kim, J. & Kim, J. (2007). Use contexts for the mobile internet: A longitudinal study monitoring actual use of mobile internet services. *International Journal of Human-Computer Interaction*, 18, 269-292.
- Ling, R. (2005). *The Mobile Connection*. Germany: Elsevier.
- Ling, R. & Yttri, B. (1999). *Nobody sits at home and waits for the telephone to ring: micro and hypercoordination through the use of the mobile telephone*. Telenor R&D Technical Report #30/99.
- Lohr, S. (2011, March). Google schools its algorithm. *The New York Times*, Retrieved January 2012 from:  
[http://www.nytimes.com/2011/03/06/weekinreview/06lohr.html?\\_r=2&pagewanted=1&hpw](http://www.nytimes.com/2011/03/06/weekinreview/06lohr.html?_r=2&pagewanted=1&hpw)
- Mandelbrot, B.B. (1977). *The Fractal Geometry of Nature*. New York: W.H. Freeman.

Manolo Project (2005). *Guidelines for integrating e-, w-, and m-learning. Project deliverables—lessons learned*. Retrieved January 2012 from:

<http://www.spinlab.vu.nl/manolo>.

Matthews, T., Pierce, J. & Tang, J. (2009). *No smartphone is an island: The impact of places, situations, and other devices on smartphone use*. IBM Technical Report #RJ10452.

Newark-French, C. (2011). Mobile apps put the web in their rear-view mirror. Retrieved June 2010 from: <http://blog.flurry.com>.

Newark-French, C. (2012). Mobile app usage further dominates web, spurred by Facebook. Retrieved Jan 2012 from: <http://blog.flurry.com/bid/80241/Mobile-App-Usage-Further-Dominates-Web-Spurred-by-Facebook>.

Nielsen, J. (1993). *Usability Engineering*. Boston, MA: Academic Press.

Nielsen, J. (2009). Mobile usability. Retrieved December 2010 from: <http://www.useit.com/alertbox/mobile-usability.html>.

Nielsen, J. (2011). Mobile usability update. Retrieved January 2012 from: <http://www.useit.com/alertbox/mobile-usability.html>.

Nylander, S., Lundquist, T. & Brannstrom, A. (2009). At home and with computer access – why and where people use cell phones to access the internet. In *Human Factors in Computing Systems: Proceedings of CHI 2009*, 1639-1642. New York: ACM Press.

- Obendorf, H., Weinreich, H., Herder, E. & Mayer, M. (2007). Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. In *Human Factors in Computing Systems: Proceedings of CHI 2007*, 597-606. New York: ACM Press.
- O'Reilly, T. & Batelle, J. (2009). Web squared: Web 2.0 five years on. White paper delivered at Web 2.0 Summit. Retrieved Jan 12 from:  
[http://gossgrrove.com/sites/default/files/web2009\\_websquared-whitepaper.pdf](http://gossgrrove.com/sites/default/files/web2009_websquared-whitepaper.pdf).
- Oulasvirta, A., Tamminen, S., Roto, V. & Kuorelahti, J. (2005). Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI. In *Human Factors in Computing Systems Proceedings of CHI 2005*, 919-928. New York: ACM Press.
- Oulasvirta, A., Petit, R., Raento, M. & Tiitta, A. (2007). Interpreting and Acting on Mobile Awareness Cues, *Human-Computer Interaction*, 22, 97-135.
- Oulasvirta, Rattenbury, Ma & Raita (2011). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, Published online: 16 June 2011.
- Palen, L. & Salzman, M. (2002). Beyond the handset: Designing for wireless communications usability. *ACM Transactions o Computer-Human Interaction*, 9, 125-151.
- Perry, M., O'hara, K., Sellen, A., Brown, B. & Harper, R. (2005). Dealing with mobility: Understanding access anytime, anywhere. *ACM Transactions on Computer-Human Interaction*, 8(4), 323-347.

Pew Research Center. (2010). Internet & American Life Project. Retrieved from:

<http://pewinternet.org>.

Pirolli, P. & Card, S.K. (1995). Information foraging. *Psychological Review*, 106, 643-675.

Pirolli, P. & Fu, W. (2003). Snif-act: A model of information foraging on the world wide web. *Proceedings of UM 2003*.

Pitkow, J.E. (1997). *Characterizing world wide web ecologies*. Unpublished Dissertation. Department of Computer Science, Georgia Institute of Technology.

Pitkow, J.E. (1998). Summary of WWW characterizations. *The Web Journal*, 2(1), 3-13.

Pitkow, J.E., Kehoe, C.J., Rogers, J. (1998). Gvu's 9<sup>th</sup> WWW user survey. Georgia Institute of Technology, Atlanta, GA. Retrieved Jan 2012 from:  
[http://www.cc.gatech.edu/gvu/user\\_surveys/survey-1998-04/](http://www.cc.gatech.edu/gvu/user_surveys/survey-1998-04/).

Purcell, K. & Brady, K. (1965). Adaptation to the invasion of privacy: monitoring behavior with a miniature radio transmitter. *Merrill-Palmer Quarterly*, 12, 242-254.

Raento, M., Oulasvirta, A., Petit, R. & Toivonen, H. (2005). ContextPhone: A prototyping platform for context-aware mobile applications. *IEEE Pervasive Computing Special Issue on Smartphone*, 4(2), 51-59.

Raento, M., Oulasvirta, A., & Eagle, N. (2009) Smartphones: An emerging tool for social scientists, *Sociological Methods Research*, 37:3, 426-454.

- Rose, D.E. & Levinson, D. (2004). Understanding user goals in web search. In *Proceedings of the World Wide Web Conference, WWW 2004*. 13-19. Banff, Alberta, Canada: ACM Press.
- Roto & Oulasvirta (2005). Need for non-visual feedback with long response times in mobile HCI. In *Proceedings of the World Wide Web Conference WWW 2005*, 775-781. Banff, Alberta, Canada: ACM Press.
- Schneiderman, B. (2000). Universal usability. *Communications of the ACM*, 43, 85-91.
- Schneiderman, B. & Plaisant, C. (2010) *Designing the User Interface: Fifth Edition*. Reading, MA: Addison-Wesley.
- Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- Shepard, C.W., Rahmati, A., Tossell, C.C., Kortum, P.T. & Zhong, L. (2010). *LiveLab: Measuring Wireless Networks and Smartphone Users in the Field*. *ACM Sigmetrics Performance Evaluation Review*, 36, 15-20.
- Simon, H.A. (1959). *Models of Man Social and Rational*. New York: Wiley.
- Simon, H.A. (1971). Designing Organizations for an Information-Rich World. In Greenberger, M. (Ed.). *Computers, Communications, and the Public Interest*. Baltimore, MD: The Johns Hopkins Press.

- Sohn, T., Li, K.A., Griswold, W.G. & Hollan, J.D. (2008). A diary study of mobile information needs. In *Human Factors in Computing Systems: Proceedings of CHI 2008*, 433-442. New York: ACM Press.
- Soskin & John (1963). The study of spontaneous talk. In Barker, R.G. (Eds.). *The stream of behavior*. New York: Appleton-Century-Crofts.
- Sugden, D. (2005). Exploring the physical frontiers of e-learning. The use of personal digital assistants (PDAs) in outreach centres. *ALT-C 2005 Research Proceedings*, September 6-8, University of Manchester.
- Sykes, R.E. (1978). Toward a theory of observer effect in systematic field observation. *Human Organization*, 37, 148-156.
- Tamminen, S., Oulasvirta, A., Toiskallio, K. & Kankainen, A. (2004). Understanding mobile contexts, *Personal & Ubiquitous Computing*, 8, 135-143.
- Tauscher, L. & Greenberg, S. (1997). How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies*, 47, 97-137.
- Taylor, C.A., Samuels, N. & Ramey, J.A. (2009). Always on: A framework for understanding personal mobile web motivations, behaviors, and contexts of use. *International Journal of Mobile Human Computer Interaction*, 1(4), 24-41.
- Tossell, C.C., Kortum, P., Shepard, C., Rahmati, A. & Zhong, L. (2010). Assessing the performance of common tasks on handheld mobile computers, In *Proceedings of the*

- Human Factors and Ergonomics Society 54<sup>th</sup> Annual Meeting*, 624-628. Santa Monica, CA: Human Factors and Ergonomics Society.
- Tossell, C.C., Schvaneveldt, R.W. & Branaghan, R.J. (2010). Targeting knowledge structures: A new rating method to elicit the relatedness of concepts. *Journal of Cognitive Technology*, 15(2), 11-19.
- Tossell, C.C., Kortum, P., Shepard, C., Barg-Walkow, L.H., Rahmati, A. & Zhong, L. (2012). A longitudinal study of emoticon use in text messaging from smartphones. *Computers in Human Behavior*, 28, 659-663.
- United Nations Telecommunications Agency (2010). Retrieved from: <http://www.itu.int>
- Van Welie & de Ridder (2001). Designing for mobile devices: A context-oriented approach. Retrieved Jan 2011 from: <http://www.satama.com>.
- Wang, Z., Lin, X., Zhong, L. & Chishtie, M. (2011). Why are web browsers slow on smartphones? In *Proceedings of the ACM International Workshop on Mobile Computing Systems and Applications*. 1-6. San Francisco, CA: ACM Press.
- Weick, K.E. (1968). Systematic observational methods. In G. Lindzey & E. Aronson (Eds.). *Handbook of social psychology, Volume 2*. Menlo Park, CA: Addison-Wesley.
- Weinreich, H., Obendorf, H., Herder, E. & Mayer, M. (2008). Not quite the average: An empirical study of web use. *ACM Transactions on the Web (TWEB)*, 2(1), 1-26.
- Weiser, M. (1991). The computer of the 21<sup>st</sup> century, *Scientific American*, 265, 66-75.



Weiss, S. (2002). *Handheld Usability*. Chichester, UK: John Wiley.

White, R. & Drucker, S. (2007). Investigating Behavioral Variability in Web Search. *In Proceedings of the World Wide Web Conference WWW 2007*. Banff, Alberta, Canada: ACM Press.

Wickens, C.D. & Hollands, J.G. (2000). *Engineering psychology and human performance* (3<sup>rd</sup> Ed.). New Jersey: Prentice Hall.

Witkin, H.A., Moore, C.A., Goodenough, D.R. & Cox, P.W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, 47, 1-64.

Zipf, G.K. (1972). *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley. Original publication (1949): Hafner Publishing Co.

**Disclaimer:** The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.